The Dissertation Committee for Melissa Sue Simper certifies that this is the approved version of the following dissertation:

Progress Toward a Combined Bacterial and Viral Gene Delivery System for Mammalian Cells

Committee:

Jaquelin P. Dudley, Supervisor

Henry R. Bose, Jr.

George Georgiou

Brent Iverson

Shelley Payne

Philip W. Tucker

Progress Toward a Combined Bacterial and Viral Gene Delivery System for Mammalian Cells

by

Melissa Sue Simper, B.S.

Dissertation

Presented to the Faculty of the Graduate School of

the University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August, 2004

Dedication

To

My parents, Herbert and Jeannette,

and

My husband, Dwayne Simper,

for their constant support

Acknowledgements

I would first like to think my advisor, Dr. Jaquelin P. Dudley, for her support during this work. Her constant encouragement and assistance provided me with the training necessary to accomplish great things. I also would like to thank the rest of my committee members: Dr. Henry Bose, Jr., Dr. George Georgiou, Dr. Brent Iverson, Dr. Shelley Payne, and Dr. Philip Tucker for their helpful advice for both my research and my dissertation.

I could not have made it here without the help of the Dudley lab members, both past and present: Jenny Mertz, Jin Seo, Sanchita Bhadra, Urmila Maitra, Mary Lozano, Dana Broussard, Quan Zhu, Farah Mustafa, and Hong Cui. I thank them all for their help and friendship, with special thanks going to Jenny Mertz and Dana Broussard for all their support both in and outside the lab.

I also would like to thank the Payne, Gottlieb, Bose, Johnson, and Tucker laboratories for sharing their reagents and constructs, and other special thanks goes to Andy Liss as well as the Payne lab for all of their helpful advice and friendships.

Finally, I would like to thank my parents, Herbert and Jeannette Mann, and my husband Dwayne Simper for all of their love and support as well as their patience during the completion of my dissertation.

Progress Toward a Combined Bacterial and Viral Gene Delivery System for Mammalian Cells

Publication No._____

Melissa Sue Simper, Ph.D.

The University of Texas at Austin, 2004

Supervisor: Jaquelin P. Dudley

Gene therapy requires efficient delivery and expression of genes in specific target cells, but most current gene therapy vectors fail to accomplish this. The goal of these experiments was development of a gene therapy system for breast cancer. The combination of viral and bacterial vectors would create one gene therapy system that achieves these goals and consists of (i) highly infectious, attenuated bacteria containing a retroviral vector that would be efficiently delivered to lymphoid cells in the gut and (ii) a mouse mammary tumor virus (MMTV)-based vector capable of trafficking from lymphocytes to mammary epithelial cells where target gene expression can occur.

In initial studies, development of a bacterial delivery system that traffics through M cells and targets lymphocytes was attempted using *Shigella*, *Yersinia*, and *Escherichia coli.* A *Shigella asd* mutant was unable to efficiently deliver genes to lymphocytes, while gene delivery by the *Yersinia* type III secretion

system caused toxicity to eukaryotic cells. *E. coli* expressing invasin, which binds β1 integrins on lymphocytes, was still unable to specifically target lymphoid cells.

In attempts to develop a retroviral gene therapy vector, an MMTV hybrid provirus was tagged with a green fluorescent protein (GFP) to follow MMTV trafficking in the mouse. However, GFP insertion into the provirus disrupted either *cis*- or *trans*-acting elements necessary for replication. To facilitate the mapping of these elements in the MMTV genome, transposon-based mutagenesis was performed, and three MMTV mutants were identified with insertions in the *env* gene. These mutants had defective cytoplasmic export of unspliced *gag* mRNA and, thus, were unable to synthesize Gag. In coinfection experiments with MMTV *env* mutants, wild-type MMTV complemented the defect in unspliced mRNA export. A novel, doubly spliced, mRNA was identified from MMTV-infected cells, and predicted protein motifs showed similarities to other retroviral *trans*-acting proteins involved in unspliced mRNA export. These results suggested that MMTV expresses a *trans*-acting protein that allows nuclear export of intron-containing RNA and should be reclassified as a complex retrovirus.

Table of Contents

List of Tables ..... **xv**

List of Figures ..... **xvi**

1. Introduction **1**

1.1 Gene therapy **1**

1.1.1 Introduction **1**

1.1.2 Non-viral vectors **2**

1.1.3 Adenoviral vectors **3**

1.1.4 Retroviral vectors **4**

1.1.5 Bacterial delivery systems **6**

1.2 Bacteria as gene therapy vectors **8**

1.2.1 *Shigella* **8**

1.2.2 *Yersinia* **11**

1.2.3 *Escherichia coli* **14**

1.3 Mouse mammary tumor virus as a potential gene therapy vector **15**

1.3.1 History **15**

1.3.2 MMTV classification **16**

1.3.3 MMTV genome organization **17**

1.3.3.1 Virion structure **17**

1.3.3.2 MMTV genome **18**

1.3.4 MMTV proteins 22

1.3.4.1 Proteins expressed from the *gag* gene 22

1.3.4.2 Proteins expressed from the *pro* and *pol* genes 24

1.3.4.3 Proteins expressed from the *env* gene 26

1.3.4.4 Protein expressed from the *sag* gene 27

1.3.5 MMTV replication cycle 28

1.3.5.1 Attachment 28

1.3.5.2 Penetration/Uncoating 29

1.3.5.3 Reverse transcription/Integration 30

1.3.5.4 Proviral expression 33

1.3.5.5 Viral assembly 35

1.3.6 MMTV life cycle 35

1.4 RNA export 39

1.4.1 Cellular export 39

1.4.2 Complex retroviral RNA export 45

1.4.3 Simple retroviral RNA export 50

1.5 Rationale for this study 53

2. Materials and Methods 55

2.1 Bacterial strains 55

2.2 Transformation methods 56

2.2.1 *Shigella*.....56

2.2.2 *Escherichia coli*.....57

2.2.2.1 Calcium chloride.....57

2.2.2.2 Electroporation.....57

2.3 Bacterial infections.....58

2.3.1 *Shigella* infection.....58

2.3.2 *Yersinia* infection.....60

2.3.3 *E. coli* infection.....61

2.3.3.1 Adherent cells.....61

2.3.3.2 Suspension cells.....62

2.4 Cell lines.....63

2.5 Plasmids.....64

2.6 Plasmid preparations.....71

2.6.1 Small-scale alkaline lysis.....71

2.6.2 Large-scale alkaline lysis.....72

2.7 Transfections.....74

2.7.1 Superfect transfection of suspension cells.....74

2.7.2 Superfect transfection of adherent cells.....75

2.7.3 DMRIE-C transfection of adherent cells.....75

2.7.4 Electroporation.....76

2.8 Transposon mutagenesis 77

2.9 Fractionation 78

2.9.1 DNA fractionation 78

2.9.2 RNA fractionation 79

2.10 RNA preparation 80

2.11 Polymerase chain reaction (PCR) 81

2.11.1 PCR 81

2.11.2 Reverse-transcriptase PCR (RT-PCR) 82

2.12 Fluorescence-activated cell sorting (FACS) 83

2.13 Reporter gene assays 84

2.14 Immunoblot analysis 85

2.15 Site-directed mutagenesis 86

3. Results 87

3.1 Development of a bacterial delivery system 87

3.1.1 *Shigella* infections 87

3.1.2 *Yersinia* infections 93

3.1.3 *E. coli* infections 104

3.2 Construction and characterization of a GFP-tagged MMTV 112

3.2.1 Analysis of GFP expression from the tagged MMTVs 112

3.2.2 Expression of viral transcripts from GFP-tagged MMTVs 117

3.2.3 Construction and characterization of a GFP-tagged MMTV with a mutated splice acceptor 121

3.2.4 Construction and characterization of a GFP-tagged MMTV containing splicing regulatory sequences 125

3.2.5 Detection of the capsid protein from the GFP-tagged MMTV mutants 128

3.2.6 Detection of unspliced mRNA in the cytoplasm of a GFP-tagged MMTV mutant 130

3.3 Development and characterization of transposon-mutagenized MMTVs 132

3.3.1 Production of transposon-mutagenized MMTVs 132

3.3.2 Development and characterization of stably integrated transposon-mutagenized MMTVs 135

3.3.3 Identification of MMTV genes necessary for trafficking to the mammary gland 141

3.3.4 Identification of mutants unable to export unspliced mRNA to the cytoplasm 145

3.4 Identification of *cis*- or *trans*-acting elements located in the MMTV provirus 150

3.4.1 Construction of a CTE-dependent reporter vector to characterize MMTV mRNA export 150

3.4.2 Detection of novel MMTV mRNAs 158

3.4.3 Complementation analysis to detect an MMTV-encoded *trans*-acting factor 165

4. Discussion 170

4.1 Development of a bacterial system for gene delivery to mammalian cells 170

4.1.1 Delivery by *Shigella* to lymphoid cells is inefficient 171

4.1.2 Gene delivery by *Yersinia* causes toxicity to eukaryotic cells 173

4.1.3 *E. coli* carrying *invA* are unable to specifically target lymphoid cells 176

4.2 Development of MMTV as a gene therapy vector 178

4.2.1 GFP-tagged MMTV clones reveal a novel viral function 179

4.2.2 GFP-tagged MMTV clones fail to express a *sag* transcript 180

4.2.3 GFP-tagged MMTV clones lack Gag protein production due to a defect in viral RNA export 182

4.2.4 Transposon mutagenesis of the MMTV provirus 183

4.2.5 Identification of MMTV sequences necessary for RNA export 186

4.2.6 MMTV lacks detectable CTE activity 188

4.2.7 Identification of an MMTV *trans*-acting factor for viral RNA export 189

4.2.8 MMTV is a complex retrovirus 195

4.3 Future directions 196

Appendix 197

References **200**

Vita **226**

List of Tables

Table 2.1. Bacterial strains provided by colleagues for use in bacterial infections 55

Table 2.2. Plasmids obtained commercially or provided by colleagues 65

Table 2.3. Primers used in this study 69

Table 3.1. Plasmids utilized in bacterial infections 89

Table 3.2. Plasmids utilized in the construction and characterization of GFP-tagged MMTV 113

List of Figures

Fig. 1.1. *Shigella* invasion of epithelial cells 9

Fig. 1.2. *Yersinia* invasion through the epithelium 13

Fig. 1.3. MMTV virion structure 19

Fig. 1.4. DNA proviral synthesis by reverse transcriptase 31

Fig. 1.5. Diagram of the MMTV mRNAs 34

Fig. 1.6. Life cycle of exogenous MMTV 37

Fig. 1.7. Regulation of protein transport by the Ran GTPase cycle 41

Fig. 1.8. Proposed structure of the Rev-response element (RRE) 47

Fig. 1.9. Domain organization of *trans*-acting factors used in RNA transport 49

Fig. 1.10. Proposed structure of the constitutive transport element (CTE) 52

Fig. 3.1. Expression of GFP in Jurkat T cells 90

Fig. 3.2. Semi-quantitative PCR analysis of vector localization in the cytoplasmic and nuclear fractions 92

Fig. 3.3. Strategy for delivery of GFP plasmids to eukaryotic cells using the type III secretion system of *Yersinia* 94

Fig. 3.4. Analysis of the population of B and T cells after bacterial infection 96

Fig. 3.5. Analysis of the population of HeLa cells after infection by *Yersinia* 98

Fig. 3.6. Analysis of the population of HeLa cells after infection by *Yersinia* at different time points 101

Fig. 3.7. Expression of GFP in HeLa cells after infection by *Yersinia* at different time points 103

Fig. 3.8. Analysis of the population of A20 cells and HeLa cells after infection by *E. coli* 105

Fig. 3.9. Propidium iodide analysis of HeLa cells after infection by *E. coli* 106

Fig. 3.10. Attachment of fluorescent bacteria to human HeLa or mouse RLmale1 cells after infection by *E.coli* 108

Fig. 3.11. Attachment of fluorescent bacteria to RLmale1 mouse T cells after infection by *E. coli* measured by different washing techniques 110

Fig. 3.12. Attachment of fluorescent bacteria to EL4b mouse T cells after infection by *E. coli* measured by different washing techniques 111

Fig. 3.13. Expression of GFP in XC cells from the tagged-MMTVs 114

Fig. 3.14. Expression of GFP from the tagged-MMTVs after Dex induction in XC cells 116

Fig. 3.15. Diagram of the splice donor and splice acceptors in the GFP-tagged MMTV 119

Fig. 3.16. Detection of viral transcripts from XC cells transiently transfected with GFP-tagged MMTVs 120

Fig. 3.17. Expression of GFP in XC cells from the tagged-MMTV containing a mutated splice acceptor 122

Fig. 3.18. Detection of viral transcripts from XC cells transiently transfected with GFP-tagged MMTV containing a mutated splice acceptor 124

Fig. 3.19. RNA sequences important for recognition by the splicing machinery 126

Fig. 3.20. Detection of viral transcripts from XC cells transiently transfected with GFP-tagged MMTV containing splicing regulatory sequences 127

Fig. 3.21. Immunoblot analysis of CA expression from GFP-tagged MMTV mutants 129

Fig. 3.22. Cytoplasmic and nuclear mRNA fractionation of XC cells transiently transfected with pHYB-MTV and GFP-tagged MMTV with a mutated splice acceptor 131

Fig. 3.23. Location of primers and trimethoprim cassettes within the mutagenized MMTV provirus 133

Fig. 3.24. Locations of transposon insertions in MMTV mutants 134

Fig. 3.25. Detection of mutant MMTV constructs in stably transfected XC cells 137

Fig. 3.26. Detection of mutant MMTV constructs in stably transfected XC cells 138

Fig. 3.27. Immunoblot analysis of CA expression from XC cells stably transfected with various MMTV constructs 140

Fig. 3.28. CD4+Vβ14+ T cell deletion analysis from injected BALB/c mice 142

Fig. 3.29. Analysis of viral transcripts from mammary glands of injected BALB/c mice 144

Fig. 3.30. Immunoblot analysis of CA expression from XC cells transiently transfected with transposon-mutagenized MMTVs 146

Fig. 3.31. Expression of *gag* mRNA in XC cells transiently transfected with transposon-mutagenized MMTVs 147

Fig. 3.32. Nuclear and cytoplasmic mRNA fractionation of XC cells transiently transfected with transposon-mutagenized MMTVs 149

Fig. 3.33. Strategy to construct a CTE-dependent reporter gene 152

Fig. 3.34. Diagram of MMTV fragments used for assays of CTE activity 153

Fig. 3.35. Activity of various MMTV fragments for mRNA export in XC cells 155

Fig. 3.36. Activity of various MMTV fragments for mRNA export in 293T cells 157

Fig. 3.37. Diagram of the known and newly identified MMTV mRNAs 160

Fig. 3.38. Detection of novel viral transcripts in XC and Jurkat stable cell lines 161

Fig. 3.39. Predicted motifs of the putative MMTV-encoded protein Rem 163

Fig. 3.40. Comparison of the NLS and NES sequences of Rev-related proteins 164

Fig. 3.41. Strategy for the complementation of the *env* mutant stable cell lines with the wild-type MMTV plasmid 167

Fig. 3.42. Nuclear and cytoplasmic mRNA fractionation of transiently transfected XC stable cell lines 169

Fig. 4.1. Domain organization of *trans*-acting factors used in RNA transport 193

1. Introduction

1.1 GENE THERAPY

1.1.1 Introduction

Gene therapy is the delivery of functional genes to treat inherited or acquired diseases. Diseases ranging from genetic disorders to various types of cancer could potentially be treated through the use of gene therapy (96,116). A major difficulty for successful gene therapy treatment is the development of a gene delivery system to target specific cells. An optimum delivery system, or vector, first requires efficient gene transfer. Other factors important for gene therapy include cell or tissue-specific delivery, facile vector construction amenable to large genes, cost-effective delivery, and importantly, vector safety. All these factors are unlikely to be incorporated into one vector and, therefore, several different vectors (each having their own advantages and disadvantages) have been developed. The main vectors used today in gene therapy are viral vectors (adenoviral and retroviral vectors, most commonly) and non-viral vectors (direct injection, gene gun). More recently, bacterial delivery systems are being developed (170,240,259).

1.1.2 Non-viral vectors

Non-viral gene therapy involves the direct transfer of DNA to a cell either through direct DNA injection or uptake of DNA through a chemical complex (122,249). The two most common methods of injection include the injection of purified DNA, usually directly into tissue or through systemic injection, and the use of a gene gun, which calls for microscopic gold beads to be complexed with the DNA and injected into cells with a helium gun (159,189). The chemical complexes for non-viral gene therapy consist of liposomes and polymers. The liposomes are usually cationic lipids that form a bilayer around the DNA and then are able to associate with the plasma membrane of the cell to allow the nucleic acid to be delivered by membrane fusion or endocytosis (209). The polymers consist of DNA conjugated to polycations, such as polylysine or polyethylenimines, which also allow the DNA to be delivered to cells (127,189). These vectors all have the advantages of: (i) simplicity to use, (ii) low cost production, and (iii) most importantly, safety since no signs of toxicity have been reported and an immune response against the DNA has not been shown. However, gene delivery is inefficient using these vectors, and the DNA delivered does not integrate into the host cell genome. Therefore, expression of the gene is transient. Due to the inefficiency and short duration of gene expression, these

vectors are now being utilized as DNA vaccines until improvements can be developed (122,127,159).

1.1.3 Adenoviral vectors

Some viruses are considered suitable for gene therapy vectors since they have evolved to deliver their genomes to target cells. Adenoviruses are double-stranded linear DNA viruses that are approximately 36 kilobases (kb) in length and are currently being used as vectors in gene therapy. These vectors are highly suitable for gene delivery since they: (i) are able to incorporate ca. 38 kb of foreign DNA, (ii) infect both non-dividing and dividing cells, (iii) have a very broad host range, (iv) can be produced to high titers, and (v) have high levels of expression of the delivered gene (13,209,249). Many generations of adenoviral vectors have been developed with the first and second generations replacing different genes, which are either not necessary for viral replication or are provided by complementation in the producer cell lines. However, one problem is the appearance of replication-competent adenovirus through homologous recombination events with the viral sequences from the producer cell lines. Furthermore, adenovirus is capable of eliciting an immune response. Thus, the immune system removes cells that express the therapeutic gene together with any viral proteins, leading to transient gene expression (116,170).

To eliminate the replication-competent adenovirus and the immune response, an adenoviral vector lacking most of the viral genes was developed. These gutless vectors use the *Cre/loxP* system to delete 25 kb of the adenoviral genome, leaving only the inverted terminal repeats and the packaging signal, and allowing the incorporation of large genes of up to 32 kb. These vectors also require a helper virus that encodes proteins needed for replication and is deleted of its packaging signal to avoid contamination of the vector with replication-competent adenovirus (116,240).

The new adenoviral vectors produced today show great promise for gene therapy, but there are still disadvantages. The gutless vectors reduce the immune response, but the viral capsid still produces an immune response that makes it difficult to use the adenoviral vectors for repeated use. Finally, although a broad host range is advantageous for the vectors, the ability to infect most epithelium-derived cell types does not allow for control of gene delivery when using systemic methods of introduction (13,212,240,259).

1.1.4 Retroviral vectors

Retroviruses are another group of viruses that are presently being used in gene therapy. These enveloped viruses contain two copies of single-stranded RNA that are copied into a DNA intermediate that then can be stably integrated

into the host cell genome. These vectors have the following advantages: (i) ease of manipulation, (ii) ability to accommodate large genes, (iii) infection of a broad range of target cells, and (iv) the ability to integrate into the host cell genome, allowing for stable expression over long periods (114,116,170).

Retroviral vectors have been developed using both simple and complex retroviruses, most commonly, the murine leukemia virus (MLV)- and the human immunodeficiency virus type 1 (HIV-1)-based vectors, respectively. Both types of retroviral vectors delete most or all of their coding regions, leaving only the *cis*-acting elements required for RNA packaging (96,116). These vectors are unable to produce a viral particle and, therefore, cell lines or co-transfected plasmids providing the viral proteins in *trans* are required. The first generation of helper cell lines expressed all viral proteins from one construct that lacked a packaging signal. However, recombination between the helper and the gene vector could restore the packaging signal to the helper, and replication-competent viruses then would be produced. Current cell lines express viral proteins from two different constructs, a *gag-pol* construct and an *env* construct, that lack significant sequence homology with each other or the RNA to be packaged (114,165). These cell lines also allow pseudotyping of the vector with different envelope (Env) proteins since Env is expressed from only one construct and is not necessary for recognizing and packaging the RNA. The most common Env

protein used is the vesicular stomatitis virus G protein, which allows a broad host range for infection as well as production of higher virus titers (96,170).

As with adenoviral vectors, the retroviral vectors also have disadvantages. The MLV-based vectors, classified as simple retroviral vectors, are unable to infect non-dividing cells, which reduces the cell types targeted for gene therapy. However, this problem can be overcome by the use of the HIV-based vectors, or complex retroviral vectors, which are capable of infecting non-dividing cells. In addition, the titers generated using retroviral vectors are low when compared to the titers generated with DNA virus vectors (114,189). The possibility of producing replication-competent viruses through recombination events also is a major concern, although many of the current retroviral vectors have been modified to reduce the amount of sequence homology between all vectors. Finally, the integration of these retroviral vectors is a random event which can not be controlled, and insertions might inactivate important genes or unintentionally activate cellular oncogenes (96,114,249). Recent examples of leukemia induction after high titer retroviral therapy have been observed (37,80).

1.1.5 Bacterial delivery systems

Gene therapy requires a delivery system that is capable of targeting specific cells without eliciting an immune response or resulting in a pathogenic

virus by recombination. Many of the current viral vectors do not fulfill at least one of these requirements. An alternative to the viral vectors is bacteria-mediated gene transfer. Bacterial delivery has the advantages of tissue tropism as well as a cell-to-cell spread mechanism that might allow bacteria to target underlying tissue layers that are not infected by viral vectors. Bacterial plasmids also are able to accommodate large amounts of coding sequence. Therefore, bacterial vectors lack the size restriction that is necessary for the packaged viral vectors. An immune response can be avoided if the bacteria are cleared from the body soon after infection occurs, either through the use of antibiotics or self-destruct features built into the vector (149,246).

Bacterial delivery systems are still being developed, yet they all share the same basic concept of plasmid delivery. Attenuated bacteria carrying the plasmid to be delivered invade target cells through their usual route of entry, usually phagocytosis. Once inside the cells, the bacteria are released from the phagocytic vacuole into the cytoplasm, where they undergo lysis as a result of antibiotic treatment, an inducible autolytic mechanism, or by engineering bacteria that require essential components that can not be provided by the target cells (43). Lysis of the bacteria allows the release of the plasmid into the cytoplasm of the target cells, where it can be transported to the nucleus and subsequently be expressed. Such delivery systems are being developed in *Salmonella*

typhimurium and *Salmonella typhi*, *Shigella flexneri*, *Listeria monocytogenes*, and recombinant *Escherichia coli* (245,246).

1.2 BACTERIA AS GENE THERAPY VECTORS

1.2.1 *Shigella*

Shigella is a gram-negative bacterium that causes shigellosis or bacillary dysentery in humans (158,223). Only 10 to 100 bacteria are needed for an infection, which can vary from mild diarrhea to a severe form that includes an acute inflammation reaction leading to colonization and destruction of the colonic mucosa. *Shigella* is transmitted from person-to-person contact or contaminated foods, and then passes through the stomach until it reaches the intestine (199).

The lining of the intestine is made up of epithelial cells that are tightly joined and form a barrier to prevent infection by microorganisms. Therefore, *Shigella* and other pathogens must find an alternative way of infecting the epithelium (200). Within the intestine, *Shigella* interacts with membranous epithelial cells (M cells), which are specialized cells in the follicular associated epithelium that overlay lymphoid follicles such as solitary nodules and Peyer's patches (Fig. 1.1). The M cells take up the *Shigella* into vacuoles through a 'ruffling process' similar to that observed with epithelial cells. However, the vacuoles are not lysed, causing the *Shigella* to be translocated through the M cell

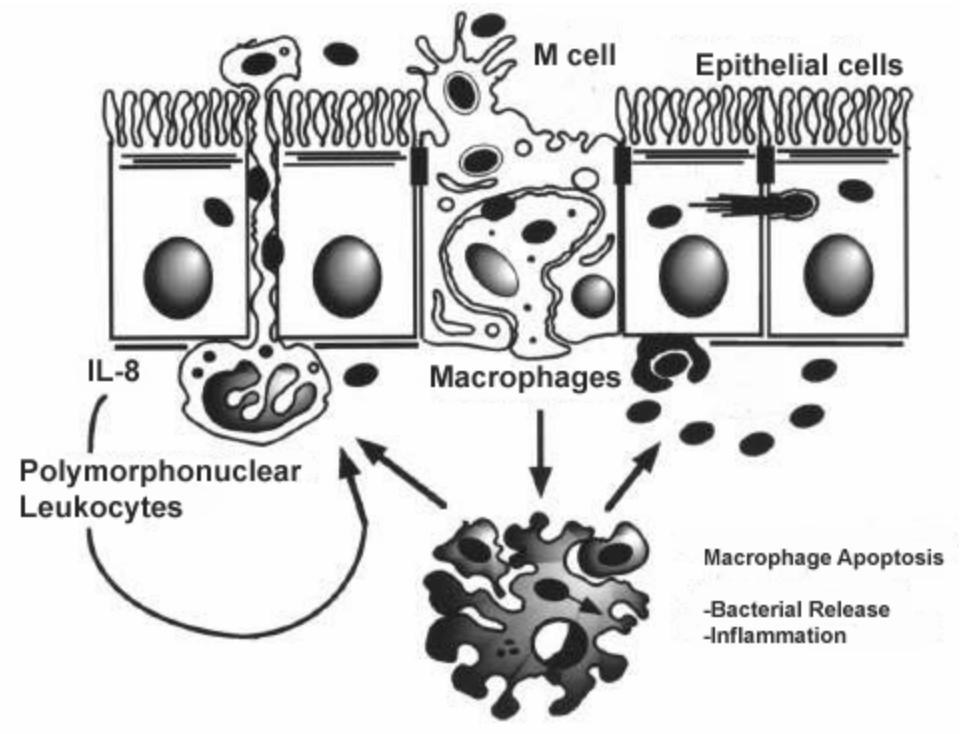


Fig. 1.1. *Shigella* invasion of epithelial cells.

Shigella crosses the epithelial barrier by translocation through M cells and subsequently infects macrophages. The infected macrophages undergo apoptosis which releases the bacteria and initiates inflammation with the release of IL-1β. Polymorphonuclear leukocyte migration to the inflammation causes destabilization of the epithelium and allows *Shigella* to invade epithelial cells through the basolateral surface. [Adapted from Sansonetti and Phalipon, 1999 (201)].

and delivered to the lymphoid follicles. *Shigella* then infects macrophages, which undergo apoptosis and release interleukin-1β, eventually leading to an inflammatory response. The inflammation attracts polymorphonuclear leukocytes (PMNs), disrupting the epithelium and allowing *Shigella* to infect these cells through their basolateral pole (198,200,201,242).

Shigella also interacts with typical epithelial cells and induces cytoskeletal rearrangements that allow the bacterium to be taken up through membrane ruffling and macropinocytosis into a vacuole. In epithelial cells, *Shigellae* lyse the vacuole, escape into the cytoplasm, and replicate. *Shigellae* move inside the cell as well as from cell to cell through actin polymerization at one pole of the bacterium, leading to further colonization of the epithelium (158,199,223).

S. flexneri is presently being studied as a bacterial delivery system since it is capable of cytoplasmic entry via vacuole lysis. To deliver plasmid DNA, several attenuated mutants have been developed (245,246). Such mutants include ones with an inability to synthesize aromatic amino acids (*aroA*) (161), deficiency in guanine nucleotide production (*guaBA*) (160), and defects in cell wall synthesis (*asd*, *dapB*) (43,207). Since the eukaryotic host cells are unable to provide the components necessary for propagation of these bacteria mutants, the bacterium will lyse and release its contents into the cytoplasm of the cell.

1.2.2 *Yersinia*

Yersinia is a gram-negative coccobacillus that is usually spread through contaminated food or water. There are three species of *Yersinia*: *Yersinia pestis* and *Yersinia pseudotuberculosis*, which are primarily pathogens of rodents, and *Yersinia enterocolitica*, which is a human pathogen. *Y. pestis* causes the plague while *Y. pseudotuberculosis* causes mesenteric adenitis and septicemia; however, *Y. enterocolitica* is able to cause a variety of infections ranging from mild diarrhea to mesenteric lymphadenitis. Although the three *Yersinia* species are pathogens of different hosts and cause different diseases, all use the same route of infection (21,42,99).

Yersinia survives at cooler temperatures and the bacterium is found in contaminated food and water in the environment, but it also must replicate once it infects a host with a higher temperature. *Yersinia* contains several pathways for host cell attachment, which are all dependent on regulation by temperature (16,99). One pathway uses the *inv* gene that encodes the protein invasin and is highly expressed at temperatures around 25^{o}C. Upon entering a human host, which raises the temperature to 37^{o}C, invasin expression is reduced, but still detectable. The other two pathways function through the chromosomally-located *ail* gene or the plasmid-encoded *yadA* gene, which are both up-regulated at 37^{o}C (21,99). Invasin allows *Yersinia* to initiate an infection by targeting β1 integrins

on M cells, resulting in invasion and translocation to the lymphoid follicles, or Peyer's patches (Fig. 1.2) (201). Subsequently, invasin is down-regulated, and the *ail* gene is up-regulated, allowing *Yersinia* to bind, but not invade, host cells. *Yersinia* then is able to replicate extracellularly and to exert effects on host cells to evade the immune system, thus causing a systemic infection. YadA assists *Yersinia* targeting of M cells as well as binding to the mucus layer, which overlays the mucosal epithelial cells, and the intestinal brush border membranes. These functions allow colonization of the intestinal tract (21,139).

Yersinia protects itself from the host immune system through the type III secretion system (TTSS) (41,124). *Yersinia* binds to macrophages and secretes proteins that: (i) impair phagocytosis, (ii) inhibit the respiratory burst, (iii) trigger apoptosis, and (iv) suppress the release of tumor necrosis factor α. *Yersinia* also binds to PMNs and secretes proteins that cause resistance to phagocytosis and inhibition of their oxidative burst. These effects prevent *Yersinia* ingestion and killing by the phagocytes and promote phagocyte killing, thereby inhibiting a specific immune response (42,66). The bacteria then can bind to other migrating cells that invade the lymph nodes, liver, or spleen to generate a systemic infection (9,44,99).

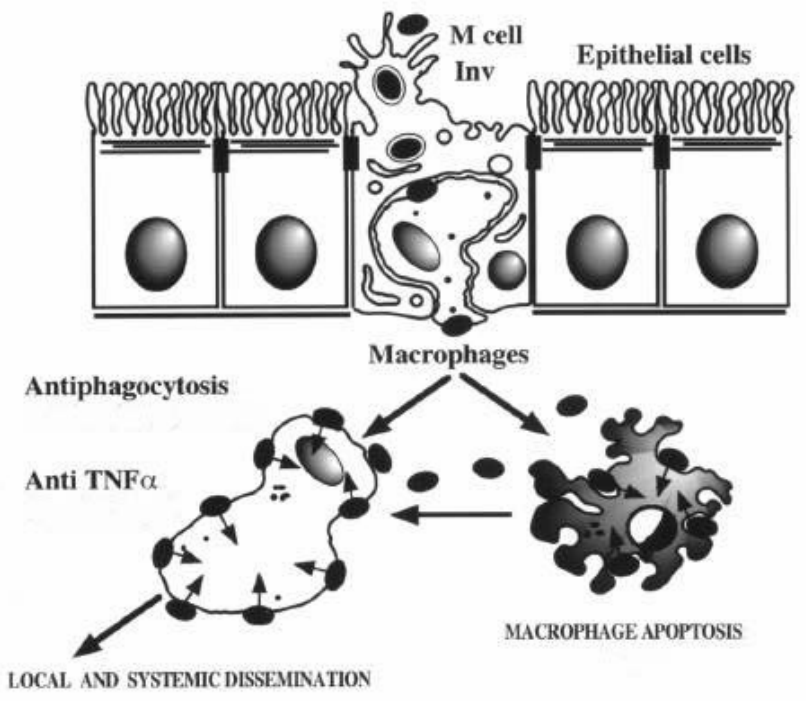


Fig. 1.2. *Yersinia* invasion through the epithelium.

Yersinia crosses the epithelial barrier by translocation through M cells. *Yersinia* begins to replicate extracellularly and evades the immune system by attaching to macrophages and injecting bacterial proteins into their cytoplasm. These proteins prevent phagocytosis or the release of TNFα in addition to inducing macrophage apoptosis. Extracellular *Yersinia* bound to macrophages and other migrating cells then are able to cause local and systemic infections. [Adapted from Sansonetti and Phalipon, 1999 (201)].

Secretion of *Yersinia* effector proteins into the cytoplasm of host cells was shown through the use of a reporter enzyme. Adenylate cyclase requires calmodulin for the production of cAMP, and calmodulin is found in eukaryotic cells, not bacteria. The catalytic domain of adenylate cyclase was fused to the N-terminal secretion and translocation domains of *Yersinia* outer membrane protein E (YopE), and *Yersinia* expressing this fusion protein were incubated with eukaryotic cells. Increase in cAMP levels indicated that translocation of the YopE-adenylate cyclase fusion protein had occurred from *Yersinia* to the eukaryotic cell, which contains calmodulin required by the adenylate cyclase to produce cAMP (210,211). These data suggest that the type III secretion system of *Yersinia* could potentially be used in gene therapy by fusing the secretion and translocation domains of YopE with a DNA-binding protein for subsequent transportation of a plasmid into a eukaryotic cell.

1.2.3 *Escherichia coli*

Escherichia coli is a gram-negative bacillus that includes both nonpathogenic and pathogenic strains. Most *E. coli* strains are nonpathogenic and are found as members of the normal intestinal flora of mammals. The related pathogenic strains of *E. coli* contain virulence factors, which allow them to cause a range of diseases including diarrhea, neonatal septicemia and meningitis, and

urinary tract infections. Pathogenic strains are usually ingested through contaminated food and include enterotoxigenic *E. coli,* enteropathogenic *E. coli,* enterohemorrhagic *E. coli,* and enteroinvasive *E. coli* (23,85,179).

E. coli differs from both *Shigella* and *Yersinia* since it targets and colonizes the intestinal epithelium instead of being translocated through the M cells. However, M cells constantly translocate antigens for presentation to cells of the immune system, and bacteria and viruses have targeted M cells to allow translocation to the Peyer's patches for subsequent proliferation in the subepithelial space (201). Benign *E. coli* strains, such as the laboratory strains K-12 or DH5α, are unable to attach to the intestinal epithelium, but might be developed as gene therapy vectors that are sampled by the M cells and translocated into Peyer's patches. The *E. coli* vectors also might be combined with other bacterial gene therapy methods, such as the TTSS, for delivery of proteins and DNA or expression of a bacterial protein for specific cell targeting.

1.3 MOUSE MAMMARY TUMOR VIRUS AS A POTENTIAL GENE THERAPY VECTOR

1.3.1 History

Mouse mammary tumor virus (MMTV) was first identified in 1933 by the Jackson Memorial Laboratory as an extra-chromosomal influence that increased

the incidence of spontaneous mammary tumors (106). Mammary tumors were shown to occur only in females and affected multiple generations. Bittner later demonstrated that an active "influence" in the milk was responsible for the high incidence of mammary tumors. This active influence was first identified as a B-type viral particle and later renamed MMTV (14,15).

1.3.2 MMTV classification

MMTV is classified as a member of the *Retroviridae* family. This family is further divided into seven genera based on sequence homology of the reverse transcriptase and according to the International Committee on Taxonomy of Viruses (178). These genera are the alpharetroviruses (e.g., avian leukosis virus), betaretroviruses (e.g., MMTV), gammaretroviruses (e.g., murine leukemia virus), deltaretroviruses (e.g., bovine leukemia virus), epsilonretroviruses (e.g., walleye dermal sarcoma virus), lentivirus (e.g., HIV-1), and spumavirus (e.g., human spumavirus). The betaretrovirus genus also contains the human endogenous retroviruses (HERVs) (235) and Jaagsiekte sheep retrovirus (129) in addition to MMTV.

MMTV also can be classified by route of transmission. Endogenous MMTVs are located in the germline of inbred mouse strains and are transmitted vertically from parent to offspring as part of the mouse genome (70,254). At least

45 endogenous MMTVs have been identified, but most are incapable of viral particle production (12). Exogenous MMTVs found in lactating mammary glands are passed from parent to offspring through the milk and induce mammary adenocarcinomas (72). Endogenous MMTVs are referred to by the nomenclature *Mtv* followed by a number (e.g., *Mtv-1*), while exogenous MMTVs are referred to by the nomenclature MMTV and the strain of mouse in which it was identified (e.g., MMTV(C3H) or C3H MMTV).

1.3.3 MMTV genome organization

1.3.3.1 Virion structure

The MMTV virion, or infectious particle, is ca. 80 to 100 nm in diameter and is distinguished from other retroviral virions by a condensed, eccentrically placed, core (31). The core of the virion contains the MMTV genomic RNA associated with several proteins and surrounded by three polypeptides. The core is enveloped and contains the viral glycoproteins observed as surface spikes. The MMTV virion first assembles as an immature core in the cytoplasm of the host cell, sometimes called intracytoplasmic A particles, and then the assembled core is directed to the plasma membrane for envelopment, release, and maturation (145). The mature, released virion protects the MMTV genome until the particle

encounters and attaches to a susceptible host cell for genome delivery and replication.

The genomic RNA is encapsidated by the nucleocapsid protein (NC) and is associated with the reverse transcriptase (RT) and the integrase (IN) proteins (Fig. 1.3). These proteins, along with the RNA, are surrounded by the capsid protein (CA), which makes up the core of the virion. The protease (PR) and matrix (MA) proteins are located outside of the core, and the MA protein is associated with a modified lipid bilayer derived from the host cell. The lipid bilayer contains two viral proteins, the transmembrane (TM) and surface (SU) proteins, which are linked together by disulfide bonds.

1.3.3.2 MMTV genome

MMTV consists of two identical, positive sense, single-stranded RNA molecules that are associated at their 5' ends (49). Each RNA genome of ca. 8.6 kb resembles cellular mRNA due to capping at the 5'end and polyadenylation of about 200 residues at the 3' end. The addition of both modifications is performed by the host cellular machinery, and the cap is added co-transcriptionally while the poly(A) sequence is added post-transcriptionally (203). Other small RNA molecules also are associated with the MMTV genome and can include ribosomal RNAs, other mRNAs, and a mixture of tRNAs. Each retrovirus has a specific

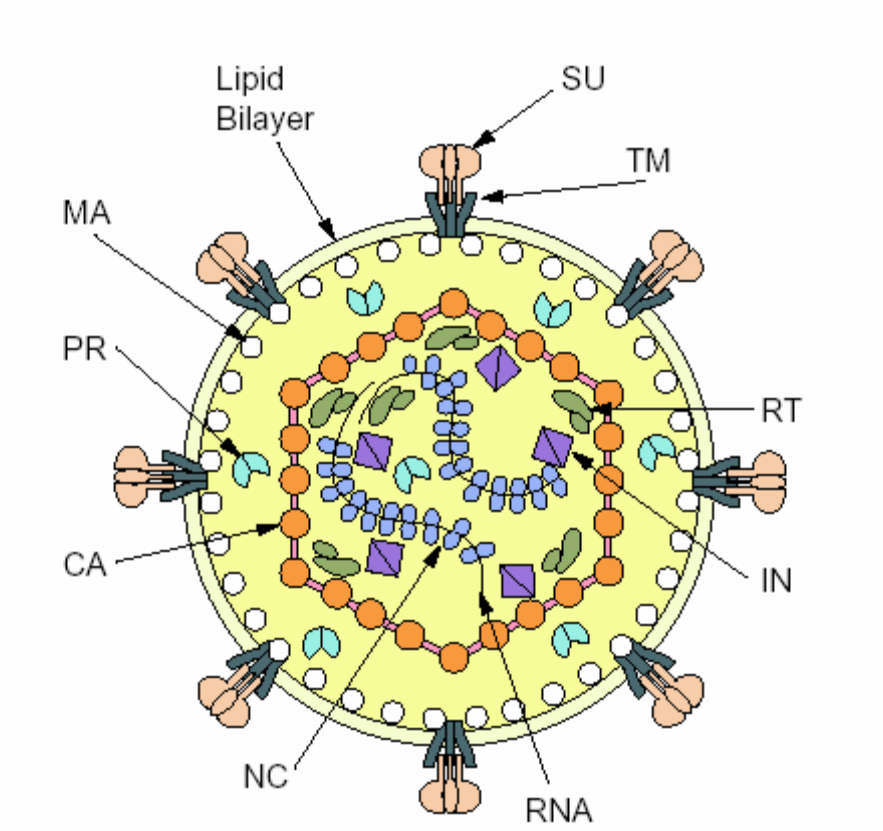


Fig. 1.3. MMTV virion structure.

The genomic RNA is encapsidated by the nucleocapsid (NC) protein and associated with reverse transcriptase (RT) and integrase (IN) proteins. The capsid (CA) protein surrounds the genomic RNA and associated proteins. The protease (PR) and matrix (MA) proteins are located outside the CA core, and the MA protein is associated with the lipid bilayer. The surface (SU) and transmembrane (TM) proteins are located in the lipid bilayer (Fig. from http://www.ncbi.nlm.nih.gov/books/bv.fcgi?call=bv.View..ShowSection&rid=rv.figgrp.2495).

tRNA that is associated with a sequence near the 5' end of the genome, which is the $tRNA^{Lys3}$ for the MMTV genome. This lysine tRNA serves as the primer for minus-strand DNA synthesis during reverse transcription (144). The MMTV RNA genome is copied into a double-stranded DNA intermediate called the provirus. The provirus includes the MMTV genes flanked by two long terminal repeats (LTR) that are identical sequences and comprised of the U3, R, and U5 regions. Integration of the provirus yields a template for transcription of viral RNA by RNA polymerase II (148,233).

The MMTV genome consists of both non-coding and coding regions. The general organization of the non-coding sequences from the 5' end of the genome is R, U5, primer binding site (PB), coding region, polypurine tract (PPT), U3, and R (see Fig. 1.4). The R region that flanks the MMTV genome is only ca. 10 base pairs in length and is the smallest R region found in retroviruses (115,237). These sequences allow the transfer of nascent DNA from one end of the genome to the other during reverse transcription. The U5 region contains unique sequences at the 5' end of the genome that directly precede the PB and becomes the 3' end of the LTR. The PB contains 18 nucleotides that bind to the $tRNA^{Lys}$ used for synthesis of minus-strand DNA (169). The PPT of the MMTV genome contains 19 A or G residues directly preceding the U3, which are not digested by RNase H during synthesis of the minus DNA strand, and are subsequently used as a primer

for the synthesis of the plus DNA strand (136,183). The final non-coding sequence of the MMTV genome is the U3 region, which contains an open reading frame (ORF) that encodes the protein Superantigen (Sag) not observed in other retroviruses. This region also encodes transcriptional regulatory signals (54).

The MMTV coding region closely resembles that of other retroviruses with the exception of the *sag* gene located in the U3 region. MMTV does not encode an oncogene like other acutely tumor-forming retroviruses and also has a longer average latency of 6 to 9 months. Therefore, MMTV is known as a non-acute, oncogenic retrovirus (72). The positions of the ORFs beginning at the 5' end of the genome are *gag*, *pro*, *pol*, *env*, and *sag* (see Fig. 1.5). The *gag* gene is translated from the full-length MMTV genome to produce the 77 kD Gag precursor, which is then cleaved to produce the MA, CA, and NC proteins. The *pro* and *pol* regions encode two polypeptides translated by ribosomal frameshifting from the full-length genome. For the *pro* region, a -1 frameshift is used to extend past the Gag precursor to produce a 110 kD polypeptide from which the PR protein is cleaved in *cis*. Two -1 frameshifts are used to produce a 160 kD polypeptide that encodes the RT and IN proteins (105). The *env* gene is translated from a singly spliced transcript that is initiated from a U3 promoter to produce the Env precursor polypeptide. This precursor is cleaved into two glycoproteins, the TM and SU proteins (92) The *env* gene also contains an

internal promoter used to initiate transcription of singly spliced *sag* mRNA, which then is translated into the Sag protein (155).

1.3.4 MMTV proteins

1.3.4.1 Proteins expressed from the *gag* gene

The *gag* gene is translated from the full-length MMTV genome to produce the three major proteins MA, CA, and NC, which represent the internal structural proteins of the virion. These proteins are found in the immature virion as the Gag precursor polypeptide, and cleavage into the final protein products does not occur until during or immediately after budding. Since the virion is assembled from the precursor polypeptide, the virions contain an equal number of the mature proteins MA, CA, and NC after cleavage of the precursor polypeptides (11).

The MA protein

The MA protein is always located at the amino (N)-terminus of the Gag polypeptide, and myristylation of this protein occurs at the glycine residue after the first methionine residue is removed (93,202). Myristate is a 14-carbon fatty acid that is co-translationally added to certain proteins that are associated with plasma membranes. The MA protein is cross-linked to phospholipids in the virion, closely associated with the plasma membrane, and seems to be a peripheral protein, which is supported by the lack of hydrophobic stretches of amino acids

found in integral membrane proteins (168). In addition to myristylation, the MA protein contains regions of basic residues that also interact with the plasma membrane. The MA protein therefore is thought to target the immature virion to the plasma membrane to promote viral assembly and budding.

The CA protein

The CA protein is the major structural protein of the viral core that surrounds the NC protein and the genomic RNA, but the structure and function of the core shell has not yet been determined (52) This protein is a hydrophobic, phosphorylated protein that contains a 20 amino acid stretch called the major homology region (MHR) found in all retroviruses, except spumaviruses. The MHR is thought to play a role in the formation of virions (219).

The NC protein

The NC protein is a small basic protein of 95 amino acids in length that is found tightly bound to the genomic RNA in the virion. The protein contains a conserved motif of cysteine and histidine residues in the sequence $CX_2CX_4HX_4C$ (CCHC) that is found either once or twice in all retroviruses, except the spumaviruses. This motif is similar to the zinc finger domains found in proteins capable of binding both zinc ions and nucleic acids. The binding of zinc to the NC protein has been shown (11). The genomic RNA is not detected in virions that do not contain the CCHC motif in the NC protein; therefore, this motif is

thought to bind to the 5' end of the genomic RNA. The NC protein also assists in the annealing of the tRNA primer to the genomic RNA as well as in the formation of the RNA dimer. The basic stretches of NC amino acids are likely to be responsible for these functions as well as participation in the assembly and budding of the virion (232).

1.3.4.2 Proteins expressed from the *pro* and *pol* genes

Like the *gag* gene, the *pro* and *pol* genes are translated from the full-length genomic RNA. However, one and two -1 frameshifts, respectively, are required to produce the polypeptides from which the mature proteins (PR, RT, and IN) are cleaved. Each of these proteins has enzymatic activity and, due to the low frequency of frameshifting, they are found in correspondingly lower levels than the structural proteins translated from the full-length MMTV genome (105).

The PR protein

The PR protein is responsible for the cleavage of the precursor polypeptides, which occurs either during budding or immediately after budding. The PR protein cleaves at mostly hydrophobic regions that are seven to eight residues in length, although the target sequences are variable (167). The sequence of the PR protein resembles the aspartic proteases that contain two aspartate residues around their active sites; however, the retroviral PR proteins are only

active as a dimer, whereas the aspartic proteases are active as monomers. This dimerization structure is thought to be important in keeping the enzyme inactive until virion assembly (208).

The RT protein

The RT protein is a highly conserved enzyme among retroviruses that is comprised of three enzymatic activities found in different domains. The enzymatic activities associated with the RT protein are the RNA-dependent or DNA-dependent DNA polymerase, which are used by the virus for double-stranded DNA synthesis from the single-stranded RNA. This polymerase domain is located in the first two-thirds of the RT protein, and polymerase activity requires a primer to initiate DNA synthesis as well as a divalent cation. Another enzymatic activity is the RNase H activity which is used by the virus to degrade the RNA portion of the RNA-DNA hybrid during reverse transcription (228).

The IN protein

The IN protein comprises the carboxy (C)-terminus of the cleaved Pol polyprotein and is less conserved compared to RT. This protein functions in the integration of the proviral DNA into the host genomic DNA. The IN protein contains a zinc finger domain that allows the protein to bind specifically to the *att* sites located at the ends of the proviral LTRs. Once bound to the proviral DNA, two nucleotides from the 3' end of the DNA are removed. The 3' OH ends attack

cellular DNA in an IN-catalyzed reaction at relatively random locations that are 6 base pairs (bp) apart on opposite strands. The IN protein then catalyzes the joining of the recessed 3' OH termini of viral DNA to the phosphates of cellular DNA. Subsequently, cellular enzymes are required to repair the chromosomal breaks and complete the viral integration event (239).

1.3.4.3 Proteins expressed from the *env* gene

The *env* gene is translated from a singly spliced transcript that is initiated from a promoter in the U3 region to produce a precursor polypeptide. The nascent precursor contains a leader sequence at the N-terminus that is bound by the signal recognition particle and directed to the membrane of the rough endoplasmic reticulum (ER). Translation of the polypeptide into the lumen of the ER continues until the polypeptide is anchored to the membrane of the ER through a hydrophobic domain at the C-terminus (53). This Env protein is transported from the ER to the plasma membrane through the Golgi apparatus, which also is the site for cleavage by a cellular protease. The two cleaved proteins then are located in the lipid envelope as a dimer held together by disulfide bonds and noncovalent interactions. The Env protein dimer probably functions to bind to receptors on host cells and allow penetration of the virus (58).

The SU protein

The SU protein is N-glycosylated after transport through the Golgi apparatus and is responsible for binding to host cell receptors. The mouse transferrin receptor 1 (mTfR1) constitutes at least one host receptor for MMTV, and is recognized by the SU N-terminus (194,260). The SU protein is the larger of the two Env proteins and is only associated with the virion through its interaction with the TM protein.

The TM protein

The TM protein spans the plasma membrane, or lipid envelope, through a hydrophobic region near the C-terminus. This protein also contains a cytoplasmic domain, or "tail", located on the internal side of the viral membrane, and this tail appears to contact the MA protein for viral assembly. The TM protein also is N-glycosylated, although not as heavily as SU (182).

1.3.4.4 Protein expressed from the *sag* gene

The Sag protein

MMTV contains an additional ORF called superantigen or Sag in the U3 region of the 3' LTR that is translated from a singly spliced transcript initiating from an internal *env* promoter (59). MMTV is the only known retrovirus that encodes this protein (2). Sag is a class II membrane protein synthesized in the ER

that undergoes proteolytic processing, possibly by furin or other protein convertases, as it is transported to the B cell membrane (51,95,185). The processed Sag then is expressed on the surface of B cells in association with the major histocompatibility complex (MHC) class II molecules where it binds to the variable portion of the beta chain (V_β) on T-cell receptor (TCR) molecules. The Sag C-terminus is variable among different MMTV strains, and each Sag reacts with specific V_β chains (253,257). Most other antigens are presented by the MHC molecules of B cells and must bind both the variable portion of the α-subunit and the V_β chain of TCR molecules. The binding of Sag to only the V_β chain therefore allows stimulation of larger numbers of T cells. This stimulation is believed to be required for amplification of the infected lymphoid cell population (90).

1.3.5 MMTV replication cycle

1.3.5.1 Attachment

MMTV, like all viruses, must attach to specific receptor molecules located on the host cell surface. This interaction occurs between MTVR1 (mTfR1) and the receptor-binding domain near the N-terminus of the Env glycoprotein SU. TfR1 is a type II, single transmembrane glycoprotein expressed by different species for iron uptake. Although multiple species express TfR1, MMTV is

specific for the rodent transferrin receptors, such as mTfR1 and rat TfR1 (194,260). Another MMTV receptor, MTVR2, mapped to mouse chromosome 19, whereas MTVR1 mapped to chromosome 16 (73). However, MTVR2 was unable to confer infection levels on virus-resistant cells comparable to mouse cells even when highly expressed, suggesting that this protein is a low-affinity receptor. Alternatively, MTVR2 may concentrate virus on the surface of the cell, allowing entry through a non-receptor-mediated pathway (73).

1.3.5.2 Penetration/Uncoating

MMTV requires host cell uptake to release genomic RNA into the cytoplasm where replication can begin. Most retroviruses undergo fusion with the plasma membrane of the host cells after attachment using a conformational change in their glycoproteins that exposes hydrophobic regions of the TM protein (57). Other retroviruses, as well as many other enveloped viruses, require an acidic environment to produce the conformational change that allows fusion. mTfR1 on the host cell normally binds transferrin and then is endocytosed into clatherin-coated pits that traffic to the acidic endosome. MMTV is dependent upon an acidic pH for membrane fusion, and acidified endosomes are required for infection. Therefore, MMTV may bind mTfR1 to allow trafficking to the acidic

endosome, where the low pH induces the conformational change necessary for membrane fusion and subsequent uncoating of the virus (194).

1.3.5.3 Reverse transcription/Integration

MMTV replication is initiated after uncoating of the virus releases the core into the cytoplasm. Once the genomic RNA is released, the proteins associated with the RNA are able to begin reverse transcription (Fig. 1.4). The RT protein initiates DNA synthesis from the $tRNA^{lys3}$ primer annealed to the PB on the 5' end of the genomic RNA (144). The genomic RNA then is copied into a minus-strand DNA, which contains sequences complementary to the R and U5 regions. The RNase H activity of RT degrades the RNA portion of the RNA-DNA hybrid, leaving single-stranded sequences called the minus-strand strong-stop DNA (187,228). The minus-strand strong-stop DNA is transferred to the 3' end of the genomic RNA where the complementary sequence of R binds to the R sequence found in the genomic RNA. Synthesis of the minus-strand DNA continues to the 5' end of the genome (the PB). RNase H degrades most of the remaining RNA genome except for the PPT and other possible oligoribonucleotide primers. The PPT or other primers then are used by RT to copy the minus-strand DNA into plus-strand DNA, and during RT polymerization, the primers are specifically cleaved and degraded by RNase H.

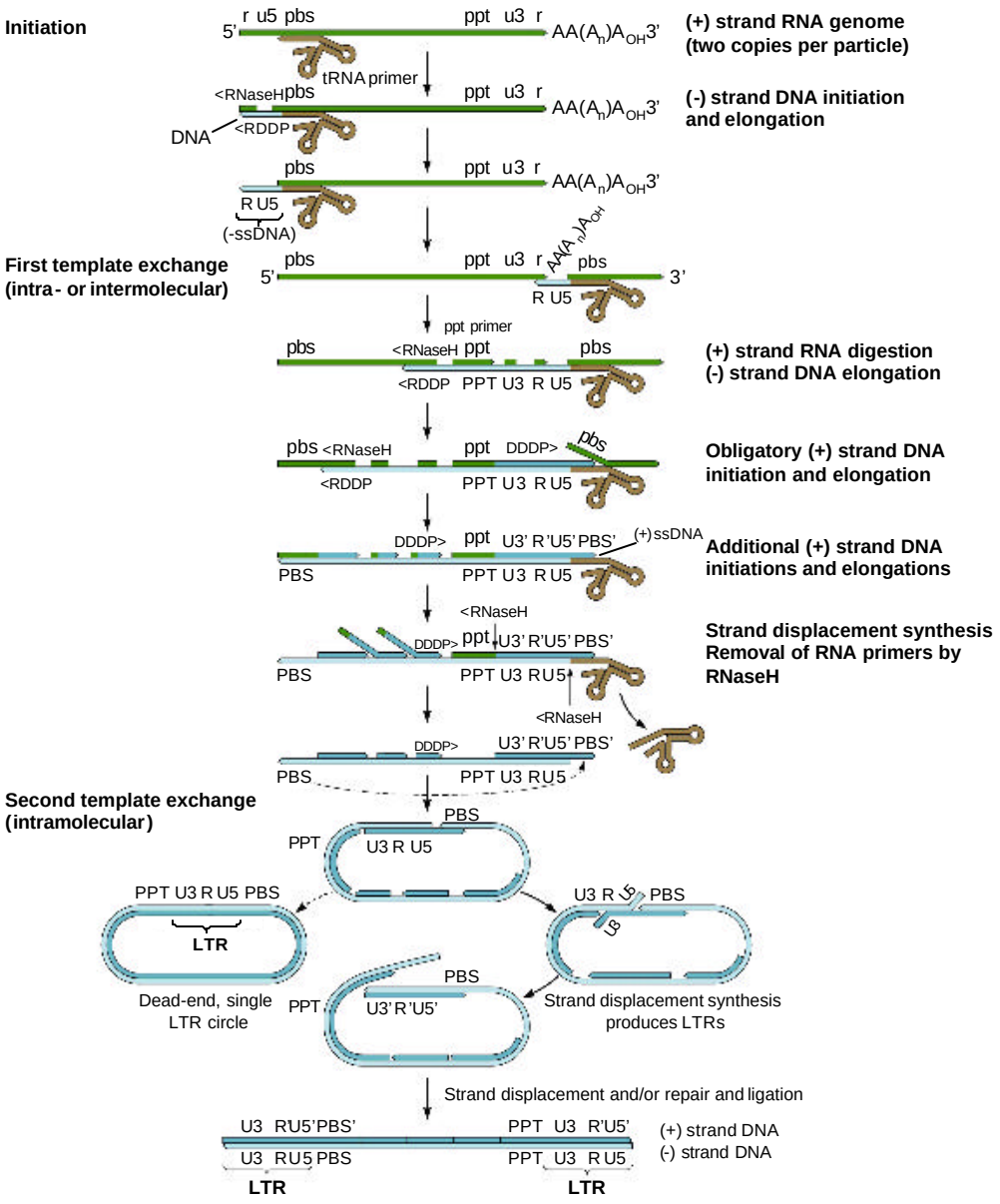


Fig. 1.4. DNA proviral synthesis by reverse transcriptase.

See text for details. [Adapted from Principles of Virology 2004 (64)]

The newly synthesized plus-strand DNA contains the U3, R, and U5 regions as well as the first 18 nucleotides of the tRNA. RNase H removes the tRNA from the RNA-DNA hybrid leaving plus-strand strong-stop DNA containing sequences complementary to the PB sequence at the 3' end of the minus-strand DNA. The plus-strand strong-stop DNA is transferred to the 3' end of the minus-strand, which is elongated by displacement synthesis. Subsequently, synthesis of the plus-strand and minus-strand DNAs are completed to form a linear double-stranded DNA, or provirus (187).

The linear double-stranded DNA synthesized by reverse transcription contains the viral genes flanked by LTRs consisting of the U3, R, and U5 regions. This DNA is still associated with many of the viral proteins, including IN, and forms a preintegration complex (243). Both LTRs contain inverted repeats at each end that are signals for IN to bind and cleave two nucleotides from the 3' ends of the viral DNA leaving a two base pair overhang and a 3' OH end (238). The complex then enters the nucleus during mitosis when the nuclear membrane is disassembled. The 3' OH ends of the viral DNA are used by IN to attack phosphodiester bonds on opposite strands of the host DNA that are 6 bp apart. The gaps created from integration are repaired and ligated by cellular enzymes, a process which also displaces the two 5' nucleotides of the viral DNA that do not

match the host DNA. The integrated proviral DNA is the template for transcription of the viral genomes (120).

1.3.5.4 Proviral expression

The MMTV provirus utilizes the host cell transcriptional and translational machinery for expression of transcripts and proteins, respectively. Reportedly, MMTV has several promoters for mRNA synthesis: two in the LTR and two in the *env* gene (Fig. 1.5). A promoter in the U3 region of the LTR is used to produce genomic and *env* mRNAs. The genomic RNA is capped and polyadenylated by host cell enzymes, and this RNA can then be used for packaging into virions or as mRNA for the translation of the Gag, Pro, and Pol polyproteins. Splicing of a portion of the genomic RNA also produces a singly spliced mRNA for translation of the Env precursor. The two U3 promoters and the two *env* promoters are all able to transcribe *sag* mRNAs. The mRNAs produced from the U3 promoters are singly spliced, whereas mRNAs synthesized from the *env* promoters may be singly spliced or unspliced (Fig. 1.5). The biological significance of multiple *sag* mRNAs remains to be determined (59,79,155).

The mRNAs then are transported to the cytoplasm for translation. The full-length genome RNA is translated by free polyribosomes into three precursor

proteins, Gag, Gag-Pro, and Gag-Pro-Pol, as previously described (105). Both *env* and *sag* mRNAs are translated on membrane-bound polysomes to allow localization of Env and Sag to the plasma membrane.

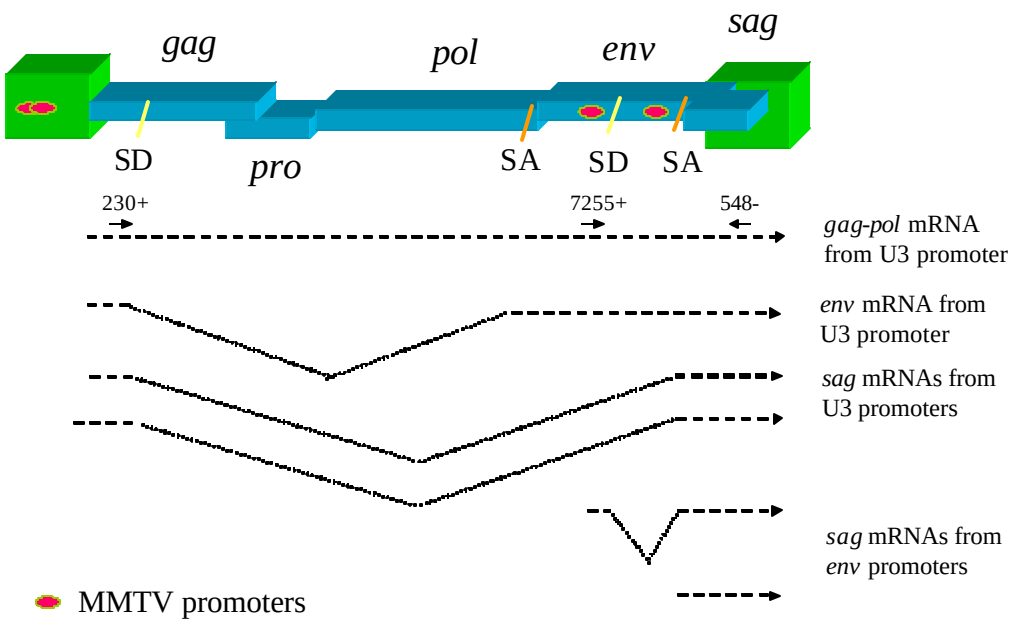


Fig. 1.5. Diagram of the MMTV mRNAs.

The MMTV provirus is shown with the LTRs as green boxes and the proteins as blue rectangles. MMTV promoters are represented by pink ovals. The mRNA transcripts expressed by MMTV are shown below the provirus as dashed lines. Splice donors (SD) and splice acceptors (SA) are indicated by black lines, and introns are shown as dotted lines in the shape of a V. Primers used to detect novel mRNAs are indicated above the *gag-pol* mRNA [Adapted from Mustafa *et al*, 2000 (155)].

1.3.5.5 Viral assembly

Although little is known about the mechanism of assembly, MMTV viral assembly begins in the cytoplasm to form an immature type A particle that then is transported to the plasma membrane for budding. It is thought that the Gag polyproteins interact with one another to initiate assembly of a virion, and the Gag-Pro and Gag-Pro-Pol polyproteins are recruited to the Gag polyprotein through targeting signals in the Gag portion of the polyprotein (145). The NC section of the Gag polyprotein appears to be responsible for packaging of the genomic RNA (123). Once the capsid associates with RNA, membrane-binding domains located in the MA portion of the Gag polyprotein target and direct the capsid to the membrane through an unknown mechanism. At the membrane, the MA portion interacts with the Env glycoproteins to allow budding. Release of virions from the plasma membrane is likely to involve a "late domain" containing the PSAP/PTAP sequences (176). During or immediately after budding, the PR protein cleaves the polyproteins in the final maturation step producing a mature virion (55).

1.3.6 MMTV life cycle

MMTV is transmitted from an infected mother to newborn pups through the milk (Fig. 1.6). The virus passes though the stomach to the intestine where it

encounters B and T cells in the gut-associated lymphoid tissue (GALT) (142,255). MMTV infects B cells in the GALT, and these B cells then express Sag on the surface in association with the MHC class II molecule. This complex allows Sag to bind specific V_β chains on the TCR molecules that stimulates a large number of T cells to produce cytokines and induces the proliferation of infected B and T cells (1,6,72). These infected B and T cells traffic to and infect dividing mammary gland cells. The infected mammary gland cells then produce MMTV virions that can be shed into the milk. A byproduct of high numbers of proviral insertions, some near cellular proto-oncogenes, in mammary cells is induction of breast cancer (72).

A number of transcriptional regulatory elements have been identified in the MMTV LTR, e.g., several negative regulatory elements (NREs), a hormone response element (HRE), and a mammary gland enhancer, that allow tissue-specific MMTV expression (26,135,151). The NRE contains binding sites for two transcription factors, special AT-rich sequence binding protein 1 (SATB1) (128) and CCAAT displacement protein (CDP) (262). SATB1 is expressed primarily in lymphocytes and suppresses MMTV transcription in these cells (34,128). CDP is a ubiquitously expressed protein that is found in undifferentiated mammary tissues, but not in lactating mammary gland. Both

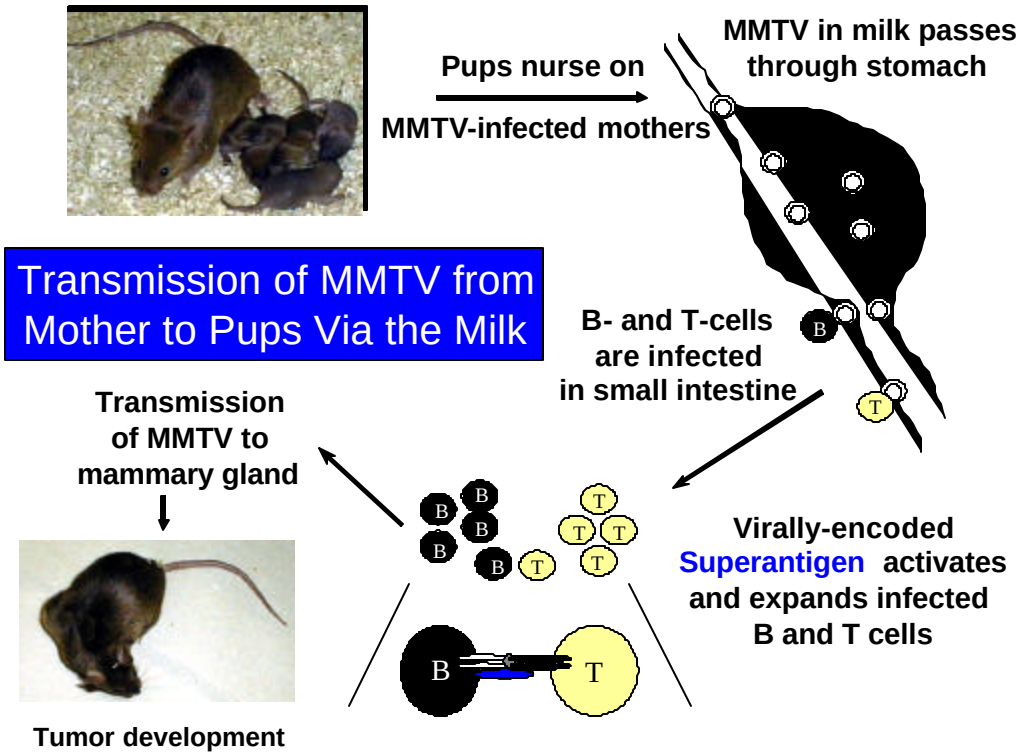


Fig. 1.6. Life cycle of exogenous MMTV.

MMTV is transmitted from an infected mother to newborn pups through the milk. The virus travels through the stomach and passes through M cells to the gut-associated lymphoid tissue (GALT). B cells in the GALT are infected by MMTV particles and express Sag on their surface in association with MHC class II molecules. Sag interacts with TCR molecules causing activation and proliferation of infected B and T cells. These infected cells traffic to the mammary gland and infect dividing mammary epithelial cells. The infected mammary gland cells then are able to produce MMTV virions that can be shed in the milk and transmitted to pups. Tumor development is likely to be a byproduct of multiple integrations sustained during peak viral replication during lactation.

SATB1 and CDP appear to suppress MMTV transcription throughout the viral life cycle until the mammary gland begins lactation (263).

The HRE and mammary gland enhancer both act as stimulators for MMTV expression in the mammary gland. The mammary gland enhancer increases MMTV RNA levels in mammary cells although the levels and identities of the bound factors are not well characterized (151). The HRE contains binding sites for steroid hormone receptors, which bind glucocorticoids and other steroid hormones (e.g., progesterone) that stimulate MMTV expression during pregnancy and lactation (86). The increased MMTV expression observed in the mammary gland ensures that a large number of MMTV virions will be produced in the milk for infection of newborn pups (151).

MMTV does not encode an oncogene, and therefore, integrations near cellular oncogenes and transcriptional disregulation of these oncogenes are required for tumorigenesis (72,236). Common integration sites, such as the *Wnt* family, the *Fgf* family, and the *Notch* family of genes, have been identified for mammary tumors induced by MMTV (77,111,188). Retroviruses randomly integrate into host DNA; thus, tumor induction occurs largely in cell types with high levels of viral replication (72).

1.4 RNA EXPORT

1.4.1 Cellular export

All retroviruses require the export of both spliced and unspliced mRNAs for replication and the use of cellular export machinery. Cellular RNA is synthesized, processed, and assembled into a ribonucleoparticle (RNP) in the nucleus, but the RNP must be transported to the cytoplasm for RNA function. Little is known about RNA export, particularly mRNA export, and therefore, comparisons have been made with other transport mechanisms. Recent evidence indicates that a number of RNA export mechanisms are dependent upon the export of proteins (156,222).

Import or export of any macromolecule occurs through the nuclear pore complex (NPC), which is located in the nuclear membrane (125,156,184). The NPC consists of several proteins called nucleoporins and contains a central channel through which these macromolecules, such as proteins, can be transported. For bi-directional transport, macromolecules must bind directly or through adaptor proteins to soluble transport receptors that target phenylalanine-glycine (FG) repeats on the nucleoporins (62,143,195,217).

Three types of transport receptors have been discovered and most belong to the karyopherin, or importin/exportin family, while the others involve a nuclear

transport factor 2 (NTF2)/p10 complex or the Tip-associated protein (TAP)/Mex67 family (244). The karyopherin family allows for transport of proteins containing nuclear localization signals (NLS) and nuclear export signals (NES), but also is involved in the transport of several classes of RNA (147,184). Since the karyopherins contain both importins and exportins, they are dependent upon the Ran GTPase cycle for directionality (Fig. 1.7). A guanine nucleotide exchange factor (RanGEF), also located in the nucleus, acts on Ran so that the predominant form in the nucleus is RanGTP. RanGTP binds to import receptors and signals the dissociation of the macromolecule from its receptor, which allows transport into the nucleus. The RanGTP-receptor complexes then return to the cytoplasm. For export of macromolecules, RanGTP in the nucleus assists in the binding of the macromolecule to the export receptor. Once transport to the cytoplasm has occurred, the cytoplasmic GTPase-activating protein and its coactivator Ran-binding protein 1 hydrolyzes RanGTP to RanGDP. RanGDP then signals the export receptor to dissociate from the macromolecule, thereby releasing it into the cytoplasm. The transport receptor NFT2/p10 imports Ran into the nucleus so that nuclear Ran is not depleted (74,76,102,118,125,156).

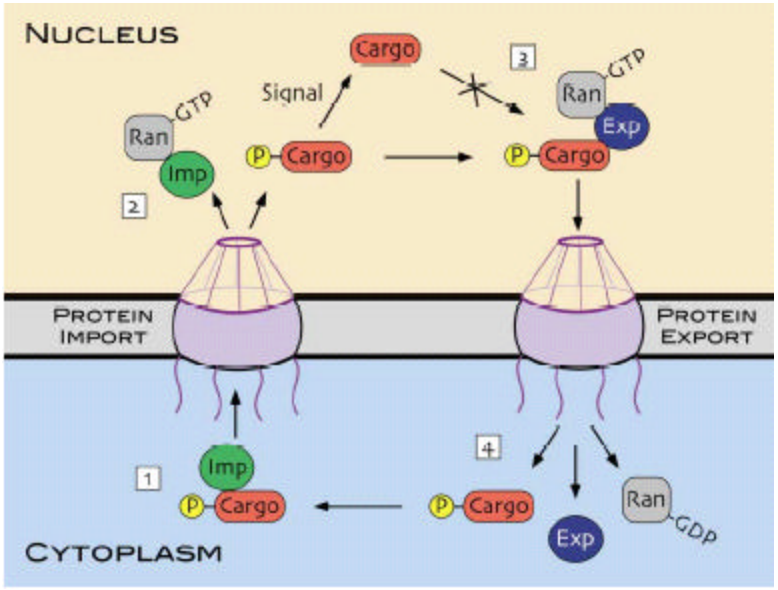


Fig. 1.7. Regulation of protein transport by the Ran GTPase cycle.

Step 1: Importin (green) binds to a cargo macromolecule (red) in the cytoplasm and targets the nuclear pore complex (NPC) (purple) for import to the nucleus. Step 2: RanGTP in the nucleus binds to the importin/cargo complex, causing the release of the cargo macromolecule. Step 3: RanGTP also binds exportin (dark blue) in the nucleus, which assists in the binding of the exportin to cargo macromolecules destined for export. This exportin/cargo complex binds the NPC to exit the nucleus. Step 4: RanGTP is hydrolyzed to RanGDP in the cytoplasm and dissociates the cargo from exportin. [Adapted from Lei and Silver, 2002 (125)].

RNA export is divided into four different classes that depend on the specific type of RNA: mRNA, rRNA, U snRNA, and tRNA. Each of these classes blocks their own export in saturating conditions, but not that of any other class, suggesting that each class uses its own export system (107,156). The karyopherin family is used for export of the rRNA, U snRNA, and tRNA classes, but an mRNA-specific exportin that allows export has not yet been discovered. Export of rRNA in the form of ribosomal subunits and 5S rRNA also is not well characterized. The export of ribosomal subunits has been difficult to analyze due to their complexity, but it has been shown that the Ran GTPase cycle is necessary and that the exportin chromosome region maintenance 1 (Crm1) protein is potentially involved (97,154). The export of 5S rRNA also likely involves Crm1 in addition to two 5S rRNA-binding proteins, TFIIIA and L5. However, interactions between these proteins and the exportin have not yet been demonstrated, and other unidentified proteins are involved since 5S rRNA mutants unable to bind to TFIIIA are exported (143,156). U snRNA export involves the 5' m^7G cap structure bound to mRNA as well as a nuclear cap-binding complex (CBC) consisting of CBP20 and CBP80. The CBC binds to the cap after addition to mRNA and then interacts with Crm1 for export of the U snRNAs (104,107). Export of tRNA is unique among the RNA export systems because tRNA binds directly to the exportin and does not require adaptor proteins.

Only tRNA that has been processed and contains a mature terminus is capable of binding to exportin-t, which then transports it to the cytoplasm (7,119).

Export of mRNA is a complicated and mysterious process, yet export appears to be linked to post-transcriptional processing of mRNAs. The pre-mRNAs transcribed in the nucleus go through 5' capping, polyadenylation, as well as splicing to form fully processed mRNAs. RNA-binding proteins differentiate between these pre-mRNAs and fully processed mRNAs and bind to the fully processed mRNAs to form RNP, which then can be exported to the cytoplasm (125,244). Cellular mechanisms enable only the fully-processed mRNAs to be exported. These mechanisms include nonsense-mediated decay, which is used to degrade mRNAs containing premature stop codons, and a retention mechanism that does not allow export of mRNAs containing introns (157,250).

Some of the first proteins identified for mRNA export were the heterogeneous nuclear ribonucleoparticle (hnRNP) proteins since these proteins could both shuttle between the nucleus and cytoplasm and bind to mRNAs (172,173). One protein, hnRNP A1, contains a sequence (M9) with both an NLS and an NES and was dependent on the karyopherin transportin. It was also shown that saturating amounts of a peptide able to bind to the M9 sequence could inhibit

export of dihydrofolate reductase (DHFR) mRNA, suggesting a role for export in a subset of the mRNAs (103,184).

The majority of mRNAs are not exported by hnRNPs; therefore, other proteins are involved in mRNA export. Mex67 was recognized as an mRNA export protein in yeast due to the nuclear accumulation of mRNA in Mex67 mutants (109,204). The human protein TAP is an orthologue of Mex67 and also appears to be involved in cellular mRNA export. TAP was shown to bind to a constitutive transport element (CTE) necessary for export of some retroviral RNAs, and excess CTE sequences blocked the export of mRNAs (78,166). The Mex67/TAP family of proteins are independent of the karyopherin family and do not require Ran GTPase, yet they still shuttle between the nucleus and cytoplasm and bind to FG repeats of the nucleoporins (38,110).

Initially, the Mex67/TAP family proteins were thought to bind directly to mRNAs, although this interaction is weak. Thus, it was suggested that mRNA export depended on adaptor proteins. The Yra1 protein was identified in yeast as a protein with binding activity for mRNA and Mex67, and it also has an orthologue (Aly/Ref) in metazoans (177,220). Aly/Ref was first identified as a cofactor for the transcription factors LEF-1 and AML and has since been shown to bind directly to TAP (33,221). Recombinant Aly/Ref in Xenopus oocytes promotes mRNA export and antibodies against Aly/Ref specifically inhibit

mRNA export, supporting a role for Aly/Ref in the export of mRNAs. Aly/Ref also provides an important link between RNA export and post-transcriptional processing. Multiprotein complexes, including Aly/Ref, form on spliced mRNAs, but not on mRNAs from cDNAs, and Aly/Ref then may direct the RNP to TAP for export (190,261).

Many mRNAs need to be exported, and another pathway may involve the export of mRNAs containing adenosine/uridine-rich elements (AREs) in their 3' untranslated regions. The RNA-binding protein HuR and phosphoproteins April and pp32 all bind to ARE-containing mRNAs, and HuR was shown to use two different export systems (30,68,244). HuR binds directly to transportin 2 through a shuttling sequence, but also uses the Crm1 pathway, possibly through interactions with April and pp32. This pathway may regulate mRNA export since, in response to heat shock, HuR loses its ability to bind transportin 2 and interacts more readily with Crm1. Because most of the heat shock mRNAs contain ARE elements, it is possible that HuR regulates the export of these mRNAs (29,67,125).

1.4.2 Complex retroviral RNA export

All retroviruses require the export of both spliced and unspliced mRNAs for replication. Unspliced mRNA serves as both the genomic RNA and the

mRNA for production of the *gag* and *pol* gene products. Complex retroviruses, such as HIV-1 and human T-cell leukemia virus type 1 (HTLV-1), also express singly spliced mRNAs in addition to completely spliced mRNAs. Cellular mechanisms only allow the export of completely spliced mRNAs; therefore, complex retroviruses must employ other methods to export their unspliced and singly spliced mRNAs (46,134,229).

Completely spliced mRNAs of complex retroviruses are exported using normal cellular pathways and are translated into regulatory proteins (46,94). One regulatory protein, HIV-1 Rev, was the first to be identified for retroviral RNA export, and proteins from other complex retroviruses (e.g., HTLV Rex1 and Rex2 and HERV-K Rec) have since been identified that also function in RNA export (84,133,174). These proteins bind to a *cis*-acting sequence found in unspliced and singly spliced mRNAs and subsequently facilitate their export into the cytoplasm. These *cis*-acting sequences are the Rev-response element (RRE) (Fig. 1.8), the Rex-response element (RxRE), and the Rec-response element (RcRE) for HIV-1, HTLV-1 and HERV-K, respectively (3,81,131,134,138,175).

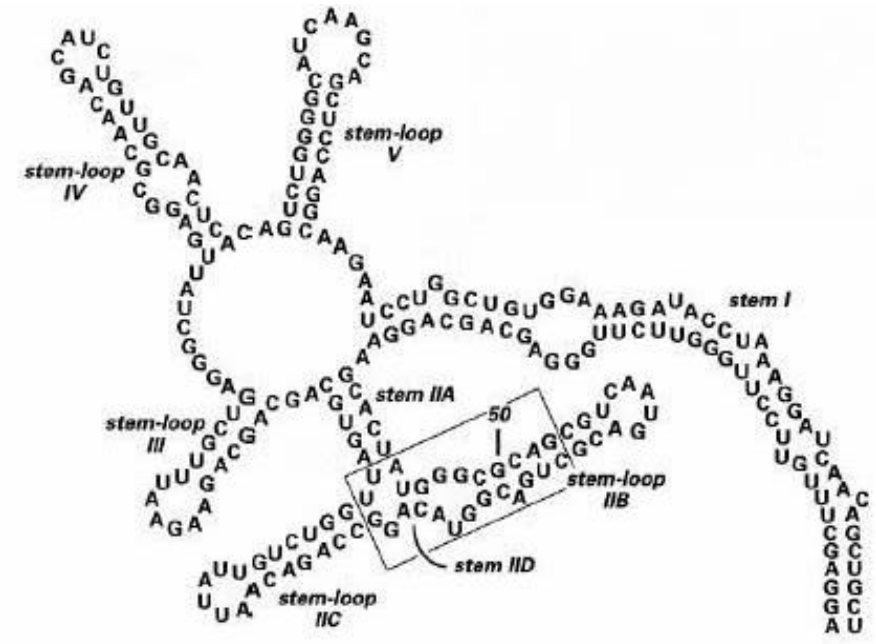


Fig. 1.8. Proposed structure of the Rev-response element (RRE).

The RRE is drawn as an RNA element containing several stem loop structures using the generally accepted nomenclature. The high-affinity binding site in the second stem loop for Rev is boxed. [Adapted from Pollard and Malim, 1998 (174)].

The Rev and Rev-related proteins must shuttle between the cytoplasm and the nucleus for RNA export. The N-terminus of each of these proteins contains a stretch of basic amino acids, mostly arginines, which acts as an NLS and nucleolar localization signal (Fig. 1.9) (39,206). Most NLSs bind to importin-α, which leads to interactions with importin-β, and this complex targets the NPC for translocation into the nucleus. However, the Rev and Rex NLSs bypass the need for importin-α and have been shown to either directly bind importin-β or be imported in the absence of importin-α, respectively. Once the proteins have been transported into the nucleus, importin-β interacts with RanGTP to dissociate the complex that delivers Rev or Rex to the nucleus (24,75,91,234). The NLS domain also functions as an RNA-binding domain (RBD) once the protein enters the nucleus and is dissociated from importin-β, and the Rev RBD binds to a high-affinity binding site in the second stem loop of the RRE (see Fig. 1.9). This bifunctional domain ensures that Rev and Rex are not bound to viral mRNAs after import into the nucleus, an event that could be harmful to viral replication (88,89,94).

Rev binds to the RRE as a monomer, but must multimerize for correct function. These multimers form through domains that allow protein-protein interactions as well as secondary binding sites on the RRE. Little is known about Rex multimer formation, although it is thought to be important for Rex function

(48,137,231). After multimer formation on the RNA, this RNP complex then is targeted for cytoplasmic export by the activation domain or the NES (Fig. 1.9). The NES contains hydrophobic amino acids, usually leucines, which are relatively evenly spaced, although different NES-containing proteins show sequence variability. The NES of these proteins then binds to Crm1 in the presence of RanGTP and is exported to the cytoplasm (63,65,113).

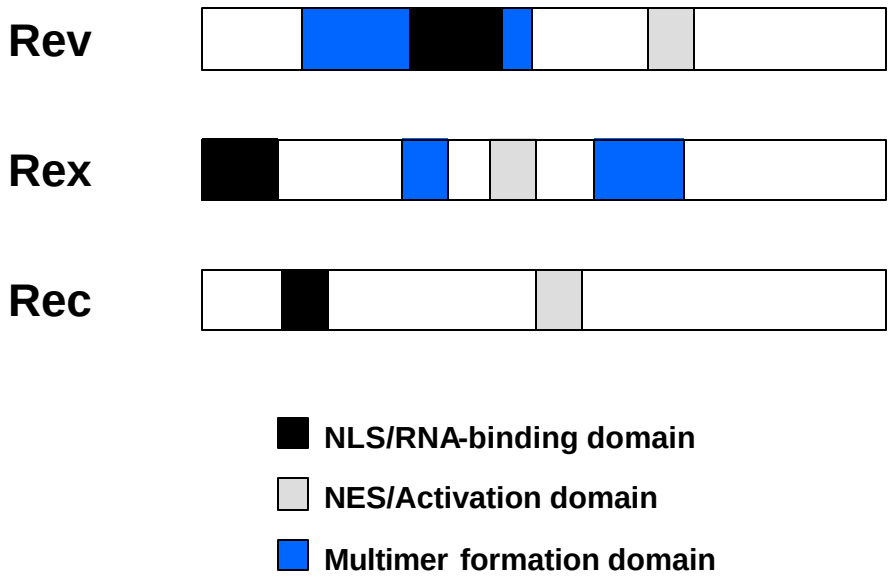


Fig. 1.9. Domain organization of *trans*-acting factors used in RNA transport.

HIV-1 Rev, HTLV-1 Rex, and HERV-K Rec are all *trans*-acting factors identified in viral RNA export. Each protein contains an NLS/RNA-binding domain near the N-terminus which functions in nuclear import and binding of the protein to RNA elements on incompletely spliced mRNAs. An NES/Activation domain is located towards the C-terminus and functions in exporting the protein/RNA complex to the cytoplasm. Rev and Rex also contain multimerization domains important for protein function.

1.4.3 Simple retroviral RNA export

Simple retroviruses, like complex retroviruses, express a spliced mRNA that uses cellular pathways for export to the cytoplasm and unspliced and partially spliced mRNAs that must use a viral mechanism for export. However, unlike the complex retroviruses, simple retroviruses do not encode a regulatory protein like Rev to export partially spliced mRNAs. Instead, such viruses contain a *cis*-acting element or CTE that interacts with cellular proteins for export of the unspliced mRNA (45,46,83).

The CTE was first discovered in Mason-Pfizer Monkey virus (MPMV) when a viral sequence was shown to substitute for the Rev and the RRE in the expression of HIV proteins from a Rev-dependent RNA (28). This element is located in the intragenic region between the *env* gene and the 3' LTR, and CTE deletion resulted in the accumulation of unspliced mRNA in the nucleus. Spliced MPMV mRNAs lacking the CTE were exported to the cytoplasm, verifying that the CTE is important for unspliced mRNA export (60,61,162). Additional simple retroviruses, e.g., the simian retrovirus type 1 and type 2 (SRV-1 and SRV-2) or intracisternal-A particle (IAP) elements, contain CTEs, and the SRV-1 CTE also has been shown to substitute for Rev and the RRE in the expression of HIV proteins (225,226,251,264). The avian sarcoma/leukemia retroviruses (ASV/ALV) contain at least one copy of a direct repeat that functions as a CTE,

but no sequence homology has been observed between the ASV/ALV elements and the CTEs (162,163,256). In addition, these elements appear to use different cellular export pathways (164).

The CTE is an RNA element that folds into a stem loop structure containing two single-stranded loops and a top loop (Fig. 1.10). The two single-stranded loops and adjacent sequences are identical, but rotated 180^{o} relative to each other, and mutations in either result in a loss of function for the MPMV and SRV-1 CTEs. These results suggest that the loops may act as binding sites for cellular protein(s) involved in RNA export (45,60,83,226). Since the CTE and Rev/RRE were interchangeable, it was expected that Crm1 would mediate RNA export through the CTE. However, competition assays determined that the CTE was unable to compete with Rev or with 5S rRNA export, but instead competed with mRNA export. Such experiments indicated that the CTE used a completely different pathway than Rev for RNA export (83,166,196).

Two cellular proteins that bound the CTE were identified as RNA helicase A and TAP. The first protein to be identified was RNA helicase A, which was thought to have a role in CTE function since (i) it promoted CTE function if overexpressed and (ii) an excess of CTE RNA localized RNA helicase A to the cytoplasm. However, antibodies against RNA helicase A were unable to prevent CTE export to the cytoplasm, and RNA helicase A also was shown to have an

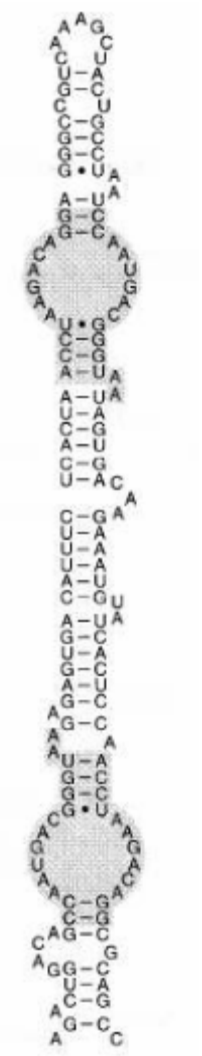


Fig. 1.10. Proposed structure of the constitutive transport element (CTE).

The CTE is drawn as an RNA element that folds into a stem loop structure containing two single-stranded loops (shaded areas) and a top loop. The shaded areas are identical but rotated 180^{o} relative to each other. These loops are the suggested binding sites for a cellular protein(s) involved in RNA export. [Adapted from Cullen, 1998 (45)].

effect on the Rev pathway. Therefore, a role for RNA helicase A in CTE function has not yet been confirmed (126,166,229). The second protein, TAP, was shown to bind to CTE RNA and also to increase cytoplasmic accumulation of CTE-containing RNA in Xenopus oocytes. TAP also plays a role in mRNA export since an excess of CTE RNA is capable of blocking mRNA export. Cellular mRNAs do not contain RNA elements like the CTE and, most likely, use adaptor proteins for TAP binding and cytoplasmic export. Therefore, the CTE may allow viral mRNAs to bypass earlier steps in the cellular mRNA export pathway and bind directly to TAP for export (45,78,83,166).

1.5 RATIONALE FOR THIS STUDY

Current gene therapy methods generally use retroviral vectors, which allow both stable integration and high-level expression of the target gene; however, targeting of these vectors to specific cells is a disadvantage that has yet to be overcome. Bacterial delivery systems target specific cells, yet gene delivery allows only transient expression. The combination of these systems in gene therapy should allow stable gene expression in specific target cells. Bacterial delivery systems able to translocate through M cells to Peyer's patches should deliver an MMTV vector to B and T cells. B and T cells carrying the MMTV vector might use Sag-specific amplification and trafficking of lymphoid cells to

deliver target genes to mammary glands. Once in the mammary gland, the MMTV vector can express a therapeutic gene, such as IL-12, to target breast cancer cells.

In this study, different bacterial systems were developed for delivery of reporter genes to eukaryotic cells. An MMTV reporter vector was constructed and characterized for use with these bacterial delivery systems to determine whether specific cell targeting and expression of the reporter gene could occur. Critical sites necessary for expression of MMTV-specific genes also were mapped by mutagenesis, and a *trans*-acting factor involved in RNA export was identified through complementation assays.

2. Materials and Methods

2.1 BACTERIAL STRAINS

The bacterial strains used in this study are listed in Table 2.1.

Table 2.1. Bacterial strains provided by colleagues for use in bacterial infections.

Strain	Characteristics	Source	Location
Shigella flexneri			
SM100	Streptomycin resistant derivative of wild type *S. flexneri*	Dr. Stefan Seliger	University of Texas Austin, TX
SM101	SM100 containing a kanamycin cassette in the *asd* gene	"	"
Yersinia pseudotuberculosis			
YP126	Wild type	Dr. Stefan Seliger	University of Texas Austin, TX
YPIII MYM	Wild type with *YadA*::Tn5 *yopH yopM yopE yopK yopO*; $Tc^R Km^R$	"	"
Escherichia coli			
DH5α invA	DH5α derivative with *Y. pseudotuberculosis invA* gene integrated into the chromosome	This study	

2.2 TRANSFORMATION METHODS

2.2.1 *Shigella*

The transformation procedure for *Shigella* was modified from Dagert and Ehrlich (47) with a prolonged incubation in calcium chloride ($CaCl_2$). A red colony from a Congo red plate was inoculated into 50 ml Luria Broth (LB) with appropriate antibiotics and incubated at 37°C with shaking until the culture reached an A_{650} = 0.2. The culture was heated to 50°C for 10 min and then was chilled on ice for at least 10 min. The *Shigella* were centrifuged in a Sorvall SS-34 rotor at 2000 x *g* for 10 min at 4°C, and the pellet resuspended in 20 ml cold 0.1 M $CaCl_2$ and incubated at 0°C for 20 to 25 min. The bacteria were centrifuged at 2000 x *g* for 10 min at 4°C, resuspended in 0.5 ml cold 0.1 M $CaCl_2$, and incubated overnight on ice at 4°C. At least 500 ng DNA was added, in a volume smaller than 10 µl, to 0.1 ml competent *Shigella* and incubated on ice for at least 30 min. The *Shigella* and DNA were subjected to heat shock at 42°C for 90 sec, and 1.5 ml LB was added prior to incubation at 37°C with shaking for 1 hr. The *Shigella* were centrifuged in a Microfuge 18 centrifuge (Beckman Coulter, Inc., Fullerton, CA) at 15000 x *g* and resuspended in a smaller volume, and the entire mixture was plated on Congo red plates with the correct antibiotics and incubated overnight at 37°C.

2.2.2 *Escherichia coli*

2.2.2.1 Calcium Chloride

E. coli were inoculated into a 5 ml LB culture without antibiotics and incubated overnight at 37°C with shaking. This overnight culture (4 ml) then was used to inoculate 500 ml LB and incubated at 37°C with shaking until an A_{600} = 0.4 to 0.45. The *E. coli* were chilled on ice for 15 to 20 min and then centrifuged in an Avanti J-E centrifuge (Beckman) with a JA-20 rotor at 6000 x *g* for 10 min at 4°C. The pellet was resuspended in 200 ml cold 0.1 M $CaCl_2$, placed into four 50 ml conical tubes, and incubated on ice for 20 min. The *E. coli* were centrifuged in an IEC Centra CL3R centrifuge at 700 x *g* for 10 min at 4°C and resuspended in 5 ml cold 0.1 M $CaCl_2$ containing 10% glycerol. DNA (30 to 50 ng) was added to 0.1 ml competent *E. coli*, mixed gently, and incubated on ice for 30 min. The bacteria and DNA were then subjected to heat shock for 90 sec at 42°C, and 0.9 ml LB added and incubated for 1 hr at 37°C with shaking. The *E. coli* were then plated onto LB plates containing the corresponding antibiotic to the plasmid resistance and incubated overnight at 37°C.

2.2.2.2 Electroporation

E. coli were inoculated into a 5 ml LB culture without antibiotics and incubated overnight at 37°C with shaking. This overnight culture (2.5 ml) was then used to inoculate 500 ml LB and incubated at 37°C with shaking until an

A_{600} = 0.5 to 0.7. The *E. coli* were chilled in an ice-water bath for 10 to 15 min and then transferred into prechilled centrifuge bottles. The *E. coli* were centrifuged in a Sorvall GS-3 rotor at 2500 x *g* for 15 min at 4°C and resuspended in 500 ml cold H_2O. The bacteria were pelleted as before and resuspended in 250 ml cold H_2O, and this step then was repeated with 125 ml H_2O, 50 ml H_2O, and finally 50 ml H_2O containing 10% glycerol. The *E. coli* were pelleted, and all but 3 to 6 ml of the supernatant was discarded. The bacteria were resuspended in 3 to 6 ml H_2O containing 10% glycerol and aliquoted for use in an electroporation. Ten ng DNA were added to 100 to 200 µl electrocompetent bacteria which were placed in a prechilled 2 mm gap cuvette. A BTX Electro Cell Manipulator 600 (BTX, Inc., San Diego, CA) was used with the settings 2.45 kilovolts and R5 (129O) to electroporate the DNA into the bacteria. One ml LB was added immediately to the bacteria and incubated at 37°C with shaking for 1 hr, and the bacteria then were plated onto LB plates, containing the antibiotics corresponding to the plasmid resistance, and incubated overnight at 37°C.

2.3 BACTERIAL INFECTIONS

2.3.1 *Shigella* infection

Shigella ssp. containing an antibiotic resistance plasmid and the construct of interest were isolated on Congo red plates supplemented with 100 µg/ml

streptomycin, 50 μg/ml kanamycin, 45 μg/ml diaminopimelic acid (DAP), and an appropriate antibiotic for the antibiotic resistance plasmid. The suspension cells were split 3 to 6 x 10^5 cells/ml so there would be 2 x 10^6 cells for each infection. On the day of the infection, 3 to 4 red colonies were isolated, and a starter culture was incubated in 1 ml LB containing the supplements indicated above for 1 hr at 37°C with shaking. Three hundred μl of the starter culture was used to inoculate 5 ml of LB containing supplements and 2.5 mM deoxycholate, and the culture was grown at 37°C with shaking to an OD_{650} = 0.5 to 0.7. A 1 ml sample of the culture was centrifuged at 15000 x *g* for 30 sec at room temperature (25°C) and resuspended in tissue culture medium (without antibiotics + DAP) to a final concentration of 2 x 10^8 colony-forming units (CFU)/100 μl. The calculation was derived from the formula OD_{650} x 8 = CFU [x 10^8] x 50 = μl tissue culture medium needed to resuspend the pellet. A larger sample of bacteria was centrifuged and resuspended with tissue culture media if needed. Isolated lymphocytes (2 x 10^6) were resuspended in 5 ml of tissue culture media containing 30 μg/ml DAP, placed in a 15 ml conical tube, and 100 μl *Shigella* (2 x 10^8) were added and mixed thoroughly. The bacteria and lymphocytes were centrifuged at 1500 x *g* for 10 min at 25°C to bring them into close contact for attachment and invasion. The cap of the 15 ml conical tube was loosened, and the tube was placed in a 5% CO_2 incubator at 37°C for 45 min. The pellet was

resuspended and incubation was continued for an additional 45 min at 37°C. The cells were washed twice with phosphate-buffered saline (PBS), and 3 ml fresh media containing 100 µg/ml gentamycin without DAP was added. The cells were incubated for 90 min at 37°C, washed once with PBS, then resuspended in 10 ml media with 50 µg/ml gentamycin. The cells were placed in a 100 mm plate and incubated for 48 hr at 37°C prior to analysis.

2.3.2 *Yersinia* infection

Adherent mammalian cells were split into 35 mm dishes the day before an infection to allow 60 to 70% confluence on the day of infection. The *Yersinia* strains were grown in brain-heart infusion (BHI) medium (Difco Laboratories, Detroit, MI) containing 2.5 mM $CaCl_2$ plus appropriate antibiotics at 26°C with shaking. To initiate the experiment, the *Yersinia* cultures (1 ml) were washed with BHI medium containing 5 mM ethyleneglycol-bis N,N'-tetra acetic acid (EGTA) and 20 mM magnesium chloride ($MgCl_2$). The *Yersinia* were then diluted 50-fold with the same media plus 3 µl/ml o-Nitrophenyl-b-D-fucoside (ONPF), incubated 1 to 2 hr at 26°C with shaking, and shifted to 37°C for 1 hr with shaking. Fresh medium containing 5 µg/ml cytochalasin D (CCD) was added to the adherent cells 30 min before the infection to stop the uptake of the bacteria. The *Yersinia* cultures were centrifuged at 15000 x *g* for 30 sec at 25°C

and resuspended with an equal volume of tissue culture media without antibiotics. Two hundred µl of *Yersinia* was added to each well of adherent cells and centrifuged in an IEC Centra GP8 at 170 x *g* for 5 min at 25°C. The mammalian cells and bacteria were then incubated in a 5% CO_2 incubator at 37°C for 30 min, after which they were washed twice with PBS and incubated an additional 45 minutes in fresh media containing 5 µg/ml CCD. After the incubation, the cells were washed twice with PBS and incubated for 1 hr at 37°C in media containing 50 µg/ml gentamycin. The cells then were washed three times with PBS, and fresh media containing 20 mg/ml gentamycin was added to the cells and incubated for 48 hr.

2.3.3 *E. coli* infection

2.3.3.1 Adherent cells

The day before the infection, adherent mammalian cells were split into 35 mm plates, and *E. coli* strains containing the test construct were inoculated into 5 ml LB with the appropriate antibiotics. On the day of the infection, the *E. coli* strains were diluted 100-fold in LB plus antibiotics and grown to mid-log phase at 37°C with shaking. The bacteria were pelleted and resuspended to a concentration of 2 x 10^8 bacteria/ml in PBS. The adherent cells were washed once with PBS, and 2 ml tissue culture media without antibiotics was added to

each well. One hundred µl *E. coli* was added to each well, and the mammalian cells and bacteria were centrifuged in an IEC Centra GP8 at 170 x *g* for 5 min at 25°C and incubated 90 min at 37°C in a 5% CO_2 incubator. The cells were washed 3 times with PBS, and 2 ml of tissue culture media plus 100 µg/ml gentamycin were added to each well and incubated for an additional 90 min at 37°C. This procedure was repeated in the presence of 20 µg/ml gentamycin and an incubation time of 48 hr prior to analysis of the cells.

2.3.3.2 Suspension cells

The day before the infection, the *E. coli* were inoculated using the protocol for infection of adherent cells into 5 ml LB with appropriate antibiotics, and the suspension cells were subcultured to approximately 5 x 10^5 cells/ml. On the day of the infection, the *E. coli* strains were diluted 10-fold in 5 ml LB with antibiotics and incubated for 2 hr at 37°C with shaking. The suspension cells were counted and resuspended to a concentration of 1 x 10^6 cells/1.5 ml tissue culture media without antibiotics and, when specified, 5 µg/ml CCD in 35 mm dishes. The *E. coli* were resuspended to a concentration of 2 x 10^8 bacteria/100 µl tissue culture media according to the *Shigella* infection procedure, and 100 µl of the bacterial suspension was added to the appropriate wells. The mammalian cells and bacteria were centrifuged at 170 x *g* for 5 min at 25°C and incubated at 37°C for 90 min. The cells were washed with PBS in 1.5 ml conical tubes by

centrifugation at 450 x *g* for 5 min at 4°C. Infected cells were resuspended in 1.5 ml tissue culture media with 50 μg/ml gentamycin and incubated 1 hr at 37°C. The cells were washed again with PBS, resuspended in 1.5 ml tissue culture media with 50 μg/ml gentamycin, and incubated overnight at 37°C prior to analysis.

2.4 CELL LINES

Jurkat human T cells were grown in RPMI medium (Invitrogen, Carlsbad, CA) supplemented with 7.5% fetal calf serum (FCS) (Invitrogen), 100 U/ml penicillin, 100 μg/ml streptomycin, and 50 μg/ml gentamycin. A20 murine B cells, LBB.11 murine B cells (obtained from Dr. Brigitte Huber, Tufts University, Boston, MA), RLmale1 murine T cells, and EL4b murine T cells (obtained from Dr. James Allison, University of California, Berkeley, CA) were maintained in RPMI medium supplemented with 10% FCS, 100 U/ml penicillin, 100 μg/ml streptomycin, 50 mg/ml gentamycin, and 5 x 10^{-5} M 2-mercaptoethanol. XC rat fibroblast cells were grown in Dulbecco's modified Eagle's medium (DMEM) (Invitrogen) supplemented with 5% FCS, 100 U/ml penicillin, 100 μg/ml streptomycin, and 50 μg/ml gentamycin. 293T human epithelial cells were maintained in DMEM supplemented with 10% FCS, 100 U/ml penicillin, 100 μg/ml streptomycin, and 50 μg/ml gentamycin. Henle human epithelial cells

(obtained from Dr. Shelley Payne, University of Texas, Austin, TX) were grown in Minimum Essential Medium (MEM) supplemented with 10% FCS, 10% tryptose phosphate broth, and 2 mM non-essential amino acids.

2.5 PLASMIDS

Expression vectors were obtained commercially, provided by colleagues, or constructed in this study (Table 2.2). Plasmids constructed in this study are described below.

The CTE-dependent luciferase vector was cloned using pRL-TK (Promega, Madison, WI) (Table 2.2). The chimeric intron was removed by complete digestion of pRL-TK with *Hin*dIII and *Nhe*I followed by blunt end formation using Klenow enzyme (New England Biolabs, Beverly, MA). The vector fragment containing the HSV TK promoter, luciferase gene, and antibiotic resistance gene was isolated by gel purification using Qiaquick gel extraction kit (Qiagen, Valencia, CA), and self-ligated to create pRL-TKΔintron (Table 2.2). The SV40 late poly(A) sequence then was replaced with the BGH poly(A) sequence in both pRL-TK and pRL-TKΔintron to increase luciferase levels. To remove the SV40 late poly(A), both luciferase vectors were linearized with *Bam*HI, followed by treatment with Klenow enzyme to create blunt ends and digestion with *Not*I. The vector pcDNA3 (Invitrogen) containing the BGH

Table 2.2. Plasmids obtained commercially or provided by colleagues.

Plasmid	Characteristics	Source	Location
pHYB-MTV	MMTV hybrid provirus in a pBR322 derivative	Dr. Harold Varmus	Memorial Sloan-Kettering Cancer Center New York, NY
penvLTR/p19ΔRI	*Hin*dIII fragment from pHYB MTV containing the *env* gene and the 3' LTR in pUC19 with a deleted *Eco*RI site	Dr. Tanya Golovkina	The Jackson Laboratory Bar Harbor, MA
pEGFP-C1	Eukaryotic GFP expression vector carrying the CMV promoter	Clontech	Palo Alto, CA
pJZ466	Eukaryotic GFP expression vector containing an IRES-GFP	Dr. Henry Bose	University of Texas Austin, TX
pCD3GFP	Eukaryotic GFP expression vector carrying the SV40 promoter	Dr. Keqin Gregg	"
pLR9	Prokaryotic GFP expression vector carrying the *dnaY* promoter	Dr. Laura Runyun-Janecky	"
pLR71	Prokaryotic GFP expression vector	Dr. Laura Runyun-Janecky	"
pCMV-EGFP	Eukaryotic GFP expression vector carrying the CMV promoter	Dr. Stefan Seliger	"
p19TLTR-EGFP	Eukaryotic GFP expression vector carrying the TBLV LTR as the promoter	This study	
pMMTV-GFP	Eukaryotic GFP expression vector carrying the MMTV LTR as the promoter	Dr. Keqin Gregg	University of Texas Austin, TX
pGEMT-Easy	Cloning system for PCR products	Promega	Madison, WI

pGL2-Promoter	Eukaryotic vector with firefly luciferase under the control of the SV40 promoter	Promega	"
pIB29MEK	*Y. pseudotuberculosis* virulence plasmid containing mutations in *yopH*, *yopM*, *yopE*, and *yopK*	Dr. Stefan Seliger	University of Texas Austin, TX
pSS351	Eukaryotic GFP expression vector carrying the CMV promoter	"	"
pSS356	*yopE* promoter region and first 49 codons fused to *lacI*	"	"
pSS357	pSS356 derivative; *yopE* promoter region and first 49 codons fused to the first 333 codons of *lacI*	"	"
pLR56	Prokaryotic vector that expresses invasion from *Y. enterocolitica*	Dr. Laura Runyun-Janecky	"
pRL-TK	Eukaryotic vector with *Renilla* luciferase under the control of the thymidine kinase promoter	Promega	
pRL-TKΔintron	pRL-TK derivative with the chimeric intron deleted	This study	
pRL-BGH	pRL-TK derivative containing the BGH poly(A) sequence	"	
pΔRL-BGH	pRL-BGH derivative with the chimeric intron deleted	"	
pTR224	pUC-based cloning vector containing the MPMV CTE	"	
2-1	penvLTR/p19ΔRI derivative containing the duplicated *env/sag* PCR fragment	"	

p19ΔSalI	pUC19 derivative with the *Sal*I site removed	“	
p19envd	p19ΔSalI derivative containing the MMTV *env* gene, the duplicated *env/sag* PCR fragment, and the MMTV 3’ LTR	“	
p19envd-GFP	p19envd derivative containing the IRES-GFP from pJZ466	“	
pBR5’	pBR322 derivative containing the MMTV 5’ LTR and the *gag* and *pol* genes	“	
pHYB-GFP	pBR322 vector backbone with the MMTV provirus containing the *env/sag* PCR fragment and the IRES-GFP	“	

poly(A) was digested with *Not*I and the blunt cutter *Pvu*II. The luciferase fragments without the SV40 late poly(A) and the BGH poly(A) fragment were isolated by gel purification, and the BGH poly(A) fragment then was ligated with each luciferase fragment to create both pRL-BGH and pΔRL-BGH (Table 2.2).

The pΔRL-BGH vectors containing different MMTV fragments and the Mason-Pfizer Monkey Virus (MPMV) constitutive transport element (CTE) were constructed by linearizing pΔRL-BGH with *Not*I or *Xho*I. The MMTV fragments were amplified by polymerase chain reactions (PCR) using primers from Table 2.3 and cloned into the vector pGEMT-Easy (Promega). Fragments were removed by digestion with *Not*I and then were cloned into the pΔRL-BGH vector that had been linearized with *Not*I. The MPMV CTE was obtained from the vector pTR224, which was provided by Dr. Farah Mustafa (Table 2.2). The MPMV CTE was isolated from pTR224 after digestion with *Xho*I and *Sal*I and cloned into the pΔRL-BGH vector that had been linearized with *Xho*I. All fragments were cloned in both the forward and reverse orientation as verified by sequencing.

Table 2.3. Primers used in this study.

Letter	Primer Name	Sequence (5'-3')
A	EGFP for	GCAAGCTGACCCTGAAGTTCATC
B	EGFP rev	GGATCTTGAAGTTCACCTTGATGC
C	C3HLTR1+	ATGCCGCGCCTGCAGCAGAAATGGT
D	gag620-	CCTCCAAATCATCCCAATCCTC
E	C3H230+	CATCACAAGAGCGGAACGGAC
F	pol4235-	CAAGGGCTATTGCTCTCTTC
G	pol4235+	GAAGAGAGCAATAGCCCTTG
H	env6337-	GGGCCCCTTTTGGAGAAAATGAGAGT
I	pol5835+	GCCACGCACTACATCATC
J	2251env-	CGTTAAGATCTGACTGCACTTGG
K	TBLVenv8509+	AGCCTTGACCAAGTGCAGTCAGATCTTAACGTG
L	3'PBR3-	CACCCTGTATATGAGTTCCC
M	C3Hpol6419+	CGAAAGACATTGGGGACC
N	C3HLTR112-	TATGTCTTTGTCTGATGGGC
O	C3Henv7461-	CCCATCCTGCTTCATACC
P	C3Henv7478+	GGTATGAAGCAGGATGGG
Q	env796-	GCTTATCTACTTGAAAGCAGCCT
R	C3HLTR1169-	GCAAGTTTACTCAAAAAATCAGCA
S	Gapdh427+	CATGTTTGTGATGGGTGTGAACCA
T	Gapdh983-	GTTGCTGTAGCCGTATTCATTGTC
U	C3Henv7255+	ATCGCCTTTAAGAAGGACGCCTTCT
V	C3HLTR548-	TACTTCTAGGCCTGTGGTCA
W	DHFR-1 FP-1	GGCGGAAACATTGGATGCGG
X	DHFR-1 RP-1	GACACTCTGTTATTACAAATCG
Y	RRE for	GAGCAGTGGGAATAGTAGG
Z	RRE rev	TCCCTAGGAGCTGTTGATCC
AA	IRES4066-	ACCTTCTGGGCATCCTTC
AB	C3HLTR408-	TCTACCTATTGGATTGGTCTTATTGG

The green fluorescent protein (GFP)-tagged MMTV vector was constructed in several steps using PCR primers that contain restriction endonuclease sites (underlined) in the 5' end for cloning purposes. The sequence that is overlapped by both the *env* and *sag* genes was amplified by PCR using the sense primer (5'-GCA GTC AGA TCT TAA CGT GCT TC-3') and the antisense primer (5'-CCG CGT CGA CTA GGT GTA GGA CAC TCT C-3'). The PCR product was then used in another reaction using the same sense primer and the antisense primer (5'-GAA GAT CTC TGG AAA ACA AGC GGC CGC GTC GAC TAG GTG TAG GAC ACT C-3') to create a fragment containing *Bgl*II sites on the flanking ends as well as internal *Sal*I and *Not*I sites. The PCR fragment was digested with *Bgl*II and cloned into the vector penvLTR/p19ΔRI that had been linearized with *Bgl*II to create the vector 2-1. To remove additional *Sal*I sites, other than the one from the PCR fragment, this vector then was digested with *Hin*dIII and cloned into the *Hin*dIII site of p19ΔSalI to create the vector p19envd. The IRES-GFP fragment was removed from the vector pJZ466 by digestion with *Not*I and *Sal*I and cloned into p19envd that had been digested with *Not*I and *Sal*I to obtain the vector p19envd-GFP. pHYB-MTV was digested with *Cla*I, and the ca. 7.5 kb fragment containing the 5' portion of MMTV was purified by agarose gel electrophoresis. This fragment was cloned into pBR322 that had been linearized with *Cla*I to generate the vector pBR5'. The 3' portion of

MMTV including the GFP-tagged *env* gene was then removed from p19envd-GFP by digestion with *BstB*I and *Xba*I and isolated following agarose gel electrophoresis. The 5.4 kb fragment was then cloned into pBR5' digested with *BstB*I and *Nhe*I to create the final vector pHYB-GFP (Table 2.2).

2.6 PLASMID PREPARATIONS

2.6.1 Small-scale alkaline lysis

Two ml bacterial cultures were centrifuged in a Microfuge 18 centrifuge (Beckman) at 15000 x *g* for 30 sec at 25°C. The pellet was resuspended in 310 µl Solution I (50 mM glucose, 25 mM Tris-HCl [pH 8.0], and 10 mM ethylenediaminetetraacetate [EDTA]) containing 400 µg/ml RNaseA, and 310 µl Solution II (0.2 N sodium hydroxide [NaOH] and 1% sodium dodecyl sulfate [SDS]) was added and gently inverted 5 to 6 times to mix and lyse the bacteria. After a 5 min incubation at 25°C, 310 µl Solution III (3 M potassium acetate [pH 4.8]) was added and mixed thoroughly. The sample was subjected to centrifugation at 15000 x *g* for 10 min at 25°C, and the supernatant was transferred to another Eppindorf tube. Isopropanol (0.6 volume) was added to the supernatant to precipitate the DNA, and the sample was pelleted at 15000 x *g* for 10 min at 25°C. The pellet was washed with 70% ethanol, air-dried for 5 to 10 min, and resuspended in 50 µl of 10 mM Tris-HCl (pH 7.4) and 0.1 mM EDTA

[TE (10:0.1)]. The concentration of the DNA was determined using a Beckman DU-6 spectrophotometer and absorbance readings at 260 nm and 280 nm.

2.6.2 Large-scale alkaline lysis

Different large-scale preparations were used depending on the nature of the vector. For pBR322-based plasmids, a colony or bacteria from a glycerol stock were used to inoculate 25 ml of LB containing 50 µg/ml appropriate antibiotics and incubated overnight at 37°C with shaking. Twenty ml of the overnight culture was added to 500 ml of LB and grown at 37°C with shaking until an absorbance reading of 1.2 at 600 nm was obtained. Chloramphenicol (Sigma, St. Louis, MO) was added to a final concentration of 170 µg/ml, and the bacteria were grown overnight at 37°C with shaking. For all other plasmids, 500 ml LB containing the appropriate antibiotic was inoculated with a colony or bacteria from a glycerol stock and grown overnight at 37°C with shaking for 12 to 16 hr.

The bacteria were centrifuged in an Avanti J-E centrifuge at 6000 x *g* for 10 min at 4°C, and the supernatant was discarded. The pellet was resuspended in 8 ml of Solution I (as described in section 2.6.1) containing 2 mg/ml lysozyme and transferred to a 50 ml conical tube. Sixteen ml of Solution II (as described in section 2.6.1) was added, and the suspension was mixed well and incubated for 5

to 10 min at 25°C. Twelve ml of cold Solution III (as described in section 2.6.1) was added and mixed by inverting the tubes several times. The mixture was incubated on ice for 10 min, and the sample was centrifuged in an IEC Centra CL3R at 2850 x *g* for 30 min at 4°C. The supernatant was strained through cheesecloth into a new 50 ml conical tube, and 70 µl RNaseA (10 mg/ml) was added. To precipitate the DNA, 0.6 volume isopropanol was added to the supernatant, mixed well, and incubated at 25°C for 20 min. The DNA then was centrifuged at 2850 x *g* for 30 min at 4°C, washed with 70% ethanol, and allowed to airdry for 5 to 10 min. The pellet was resuspended in 3.3 ml of 10 mM Tris-HCl (pH 7.4) and 10 mM EDTA [TE (10:10)], and 4.6 g CsCl was added and dissolved. Four hundred µl ethidium bromide (EtBr) (5 mg/ml) was added, and the debris (RNA and protein) was removed by centrifugation at 2850 x *g* for 20 min at 4°C. The sample was added to a Beckman Optiseal tube and subjected to centrifugation in a Beckman L7 Ultracentrifuge at 220000 x *g* for 16 to 18 hr at 20°C in an NVT 65.2 rotor. The plasmid band was collected using a 5 ml syringe and an 18 gauge needle, and the EtBr extracted 3 to 4 times using 2 volumes of G-50 buffer saturated with n-butanol. A solution of TE (10:10) was added to a final volume of 5 ml, and the DNA was precipitated by adding 2 volumes ethanol and incubating at -20°C for 30 min. The DNA was pelleted in an IEC at 2850 x *g* for 30 min at 4°C, and the pellet was resuspended in 2 ml of 10 mM Tris-HCl (pH

7.4) and 1 mM EDTA [TE (10:1)]. The sample then was dialyzed against 2 L of TE (10:1) using a Slide-A-Lyzer dialysis cassette (Pierce, Rockford, IL), with buffer changes after 1 hr and overnight incubations. The DNA was collected, ethanol precipitated, and washed with 70% ethanol before resuspending in TE (10:0.1). The concentration of the DNA was determined using a Beckman spectrophotometer and absorbance readings at 260 nm and 280 nm.

2.7 TRANSFECTIONS

2.7.1 Superfect transfection of suspension cells

Suspension cells were subcultured at approximately 3 x 10^5 cells/ml the day before the transfection. The next day the cells were counted and, for each transfection, 1.25 x 10^6 cells were plated in 2.5 ml tissue culture medium in 60 mm dishes. Two µg test DNA plus 0.5 µg co-transfectant DNA was added to 75 µl tissue culture medium without serum or antibiotics. Ten µl Superfect Reagent (Qiagen) was added and mixed with the DNA, and the mixture was incubated for 5 to 10 min at 25^{o}C. The Superfect-DNA mixture was added dropwise to the suspension cells and incubated for 48 hr.

2.7.2 Superfect transfection of adherent cells

The day before the transfection, adherent cells were counted, and approximately 8 x 10^5 cells were plated into one well of a 6-well plate in tissue culture medium. One µg test DNA plus 1 µg co-transfectant DNA was added to 75 µl tissue culture medium without serum or antibiotics. Ten µl Superfect reagent was added and mixed with the DNA, and the mixture was incubated for 5 to 10 min at 25^{o}C. The cells were washed with 1X PBS, 1 ml complete tissue culture medium was added to the Superfect-DNA mixture, and approximately 1.1 ml of the mixture was added to the cells and incubated for 2 to 3 hr at 37^{o}C in an incubator with 7.5% CO_2. The medium was removed from the cells, 2 ml fresh medium was added, and the cells incubated for an additional 48 hr.

2.7.3 DMRIE-C transfection of adherent cells

The day before the transfection, adherent cells were counted, approximately 2 x 10^5 cells were plated into one well of a 6-well plate in 2 ml tissue culture medium. Ten µl of DMRIE-C reagent (Invitrogen) were mixed with 0.5 ml tissue culture medium without serum or antibiotics for each transfection and incubated for 10 to 15 min at 25^{o}C. Five µg test DNA plus 0.5 µg co-transfectant DNA was mixed with 0.5 ml tissue culture medium without serum or antibiotics. The DMRIE-C/medium mixture was added to the

DNA/medium mixture at a 1:1 ratio containing 0.5 ml of each mixture and incubated for 30 to 45 min at 25°C. The adherent cells were washed twice with tissue culture media without supplements, and 1 ml transfection mixture was added to each well and incubated for 6 to 9 hr at 37°C in an incubator with 7.5% CO_2. The transfection was stopped by the addition of 1 ml complete tissue culture medium containing 10% FCS and incubation continued for an additional 48 hr.

2.7.4 Electroporation

Electroporations were performed using a BTX ECM600 electroporator. A20 cells were seeded at 6 x 10^5 cells/ml, and the next day, the cells were counted, and a suspension of 2.5 x 10^7 cells/ml was prepared in RPMI containing 10% FCS. Twenty µg test DNA was mixed with 400 µl (1 x 10^7 cells), and the mixture was added to a 4 mm gap cuvette and incubated for 10 min at 25°C. Cells were subjected to electroporation using 280 V, 960 µF, and R4 (72O) and, subsequently, the sample was incubated on ice for 10 min. The cells then were plated in 5 ml complete RPMI in 60 mm dishes and analyzed after 24 hr. Jurkat cells were seeded at 5 x 10^5 cells/ml the day before the electroporation. The next day the cells were counted, and a suspension of 1 x 10^7 cells/ml was prepared in RPMI containing 5% FCS. One µg of reporter DNA and 2 µg of co-transfectant

DNA were added to 400 µl (4 x 10^6 cells), and the mixture was placed into a 4 mm gap cuvette and incubated for 10 min at 25°C. Conditions for electroporation were: 270 V, 1050 mF, and R10 (720O). After electroporation, the sample was incubated at 25°C for 10 min before plating in 5 ml RPMI in a 60 mm dish. Cells were harvested 48 hr after electroporation.

2.8 TRANSPOSON MUTAGENESIS

Insertion of the trimethoprim cassette into pHYB-MTV was performed using the EZ::TN™ <DHFR-1> Insertion Kit (EPICENTRE, Madison, WI). pHYB-MTV (0.2 µg) was incubated with an equimolar amount of the EZ::TN <DHFR-1> transposon and 1 U EZ::TN transposase in a 10 µl reaction. The mixture was incubated 2 hr at 37°C, and the reaction was terminated by addition of SDS to a final concentration of 0.1% and incubation for 10 min at 70°C. One µl of the reaction was used to electroporate *E. coli* DH5α cells using a Bio-Rad Gene Pulser™ (Bio-Rad Laboratories, Hercules, CA) (1.7 kV, 200 ohms and 25 mF). SOC medium was added immediately to the cells, transferred to a tube, and incubated for 1 hr at 37°C with shaking. Clones with the insertion were selected on trimethoprim-containing plates (100 µg/ml). Individual colonies were then re-isolated on trimethoprim (100 µg/ml)/ampicillin (50 µg/ml)-containing plates to select against clones with an insertion in the ampicillin gene.

2.9 FRACTIONATION

2.9.1 DNA fractionation

Cytoplasmic and nuclear DNA fractions were prepared from tissue culture cells after transfection or infection. Cells were washed with 1X PBS and counted using a hemocytometer. For every 2 x 10^7 cells, 1.5 ml cold low-salt buffer (LSB) (10 mM sodium chloride (NaCl), 3 mM magnesium acetate, and 20 mM Tris-HCl [pH 7.4]) was added, and the cell pellet resuspended. Lysis buffer (0.5 ml) (5% sucrose and 1.2% Triton-X 100 in LSB) was added and mixed, and the nuclei were pelleted using an IEC Centra CL3R centrifuge at 700 x *g* for 10 min at 4°C. The cytoplasmic fraction was removed, and EDTA was added to a final concentration of 25 mM. The DNA from the cytoplasmic fraction was precipitated with ethanol after adjustment to 0.2 M NaCl followed by an incubation of 30 min at -20°C. The DNA was centrifuged in an Avanti J-E centrifuge for 30 min at 12100 x *g* at 4°C, resuspended in 2 ml Solution IV (75 mM NaCl, 25 mM EDTA, 20 mM Tris-HCl [pH 7.4], and 0.5% SDS) with 100 µg/ml proteinase K, and incubated for 1 hr at 37°C. To the remaining nuclear fraction, 5 ml Solution IV containing 100 µg/ml proteinase K was added. The nuclear DNA was sheared six times by passage through a 25 gauge needle six times and incubated for 1 hr at 37°C. Both fractions were extracted once with a

mixture of equilibrated phenol-chloroform (1:1; v/v) and precipitated with ethanol as previously described. The DNA was centrifuged at 12100 x *g* for 30 min at 4^{o}C, washed with 70% ethanol, and resuspended in 2 ml of TE (10:0.1).

2.9.2 RNA fractionation

The cytoplasmic and nuclear RNA fractions were obtained using the DNA fractionation protocol with modifications. All solutions were treated with 0.1% diethyl pyrocarbonate (DEPC), and 8 µl/ml polyacryl carrier was added for each ethanol precipitation. DEPC-treated Solution IV (2.5 ml) was added to both fractions, and 200 µg/ml proteinase K was added to the cytoplasmic fraction keeping the nuclear fraction at 100 µg/ml proteinase K. After the ethanol precipitation, both fractions were resuspended in 300 µl TE (10:0.1), and the DNA contaminants were removed after RNA precipitation using 3 volumes of 4 M sodium acetate (NaOAc) (pH 5.3). The RNA from both fractions then was pelleted at 15000 x *g* for 30 min at 4^{o}C and washed with DEPC-treated 70% ethanol. The RNA was resuspended in DEPC-treated H_2O, and the concentration was determined from the absorbance at 260 nm.

2.10 RNA PREPARTION

Total RNA was prepared from tissue culture cells by adding 1ml of RNA extraction buffer (4 M guanidine isothiocyanate, 25 mM sodium citrate, [pH 7.0], 0.1 M 2-mercaptoethanol, 0.5% N-laurosarcosine, 50% water-saturated phenol to adjust to pH 7.0, and 0.2 M NaOAc [pH 4.0]) directly to a 35 mm culture plate containing approximately $2x10^6$ cells. One-fifth volume of chloroform was added, and the solution was mixed and the phases were separated in a Microfuge 18 centrifuge at 15000 x *g* for 5 min at (25^oC). The upper phase was collected into another tube, three-fifths volume isopropanol was added and mixed, and then the solution was incubated for 20 min at 25^oC. The RNA was precipitated by centrifugation at 15000 x *g* for 30 min at 4^oC. The pellet was resuspended in 100 µl of DEPC-treated TE (10:0.1). The RNA was further precipitated with three volumes of 4 M NaOAc (pH 5.3) to remove DNA contaminants, and the mixture was incubated for 15 min at (25^oC). The mixture was centrifuged at 15000 x *g* for 15 min at 4^oC and washed with 70% DEPC-treated ethanol. The RNA pellet was resuspended in 100 µl DEPC-treated H_20, and the concentration was determined by absorbance at 260 nm using a Beckman spectrophotometer.

2.11 POLYMERASE CHAIN REACTION (PCR)

2.11.1 PCR

Five-fold dilutions were performed on the DNA recovered from DNA fractionation of the transfections and infections, and PCR was performed to compare the DNA levels. Five µl of each dilution was used in a 50 µl reaction using *Taq* polymerase (Invitrogen) and GFP primers using the protocol provided by Invitrogen. The PCR conditions were: denaturation at 94°C for 3 min, 25 cycles of denaturation at 94°C for 1 min, annealing at 55°C for 1 min, and extension at 72°C for 1.5 min, and a final extension step at 72°C for 10 min in a PTC-100 programmable thermal controller (MJ Research, Inc., Watertown, MA.)

Colony PCR was performed on the mutagenized clones of pHYB-MTV to localize the integration site. A bacterial clone was inoculated into 5 tubes each containing 5 µl H_2O and used for colony PCR with 5 different primer pairs (C/D, E/F, G/H, I/J, and K/L) (Table 2.3) and Klentaq polymerase (AB Peptides, St. Louis, MO) in a 20 µl reaction. The PCR conditions were: denaturation at 94°C for 3 min, 30 cycles of denaturation at 94°C for 1 min, annealing at 55°C for 1 min, and extension at 72°C for 1.5 min, and a final extension step at 72°C for 10 min.

MMTV fragments from pHYB-MTV were obtained by PCR for cloning into the pΔRL-BGH vector. Ten ng pHYB-MTV was used with PfuTurbo (Stratagene, La Jolla, CA) and different primer pairs using the conditions provided by the manufacturer in a 50 µl reaction. The PCR conditions were: denaturation at 94°C for 3 min, 30 cycles of denaturation at 94°C for 1 min, annealing at 55°C for 1.5 min, and extension at 72°C for 1.5 min.

2.11.2 Reverse-transcriptase PCR (RT-PCR)

Ten µg of RNA from tissue culture cells was diluted with DEPC-treated H_20 to a final volume of 41.5 µl. The RNA then was treated with 3 U amplification-grade DNase I (Invitrogen) and 5 U RNase inhibitor (Invitrogen) for 1 hr at 37°C. The DNase I was inactivated by addition of EDTA to a final concentration of 2.5 mM and incubation at 72°C for 15 min. Five µg of the DNase I-treated RNA was further processed to make cDNA in a reverse transcriptase (RT) reaction. The primer poly(dT_{17}) and dNTPs were added to the RNA in final concentrations of 2.5 pmol/µl and 1 mM, respectively, boiled for 5 min, and immediately placed on ice for 5 min. The RNA was reverse transcribed using 400 U Moloney murine leukemia virus reverse transcriptase, 5 U RNase inhibitor, and 10 mM dithiothreitol (DTT) in a 50 µl reaction. Two and a half µl of cDNA was used in each RT-PCR using specified primer pairs and polymerases

for each mRNA (Table 2.3). Taq polymerase was used for PCR products less than 1 kb. PCR conditions were: denaturation at 94°C for 3 min, 50 cycles of denaturation at 94°C for 1 min, annealing at 55°C for 1 min, and extension at 72°C for 1 min, and a final extension step at 72°C for 5 min in a PTC-100 programmable thermal controller. Expand Long Template polymerase (Roche, Indianapolis, IN) was used for PCR products larger than 1 kb. PCR conditions were: denaturation at 94°C for 2 min, 30 cycles of denaturation at 94°C for 20 sec, annealing at 55°C for 30 sec, and extension at 68°C for 6 min, and a final extension step at 68°C for 8 min using the same thermocycler.

2.12 FLUORESCENCE-ACTIVATED CELL SORTING (FACS)

Cells were harvested after infections or transfections and pelleted by centrifugation at 450 x *g* for 5 min at 4°C using an IEC Centra CL3R centrifuge. The cells were resuspended in 1 ml 1X PBS, and 300 µl of the cells were removed and pelleted in a Microfuge 18 centrifuge at 450 x *g* for 5 min at 25°C. The cells were resuspended in 300 µl fresh 1X PBS, transferred to Falcon polysterene round-bottom tubes (Becton Dickenson, Franklin Lakes, NJ), and analyzed for GFP using a FACSCaliber flow cytometer and CELLQuest software. Forward scatter and side scatter plots were used to determine the cell population to be selected for GFP analysis, and at least 10,000 cells were analyzed. The

percentage of GFP-positive cells was then used to determine if DNA had been delivered from bacteria to eukaryotic cells or to normalize for DNA uptake in a transfection.

2.13 REPORTER GENE ASSAYS

Luciferase assays were performed using the Dual-Luciferase Reporter Assay System (Promega) to quantitate *Renilla* luciferase and to normalize for DNA uptake with either firefly luciferase or GFP-positive cells. Briefly, cells were harvested in 1X passive lysis buffer (Promega) and subjected to two freeze-thaw cycles in a dry ice/ethanol bath. The lysates were centrifuged for 5 min at 15000 x *g* at 4°C to remove cellular debris, and the supernatant transferred to another tube. The protein concentration was determined using the Bio-Rad Protein assay by comparison to a bovine serum albumin standard. A total of 40 µg or 20 µl was used to obtain luciferase readings using a Turner TD-20e luminometer (Turner Designs, Inc., Sunnyvale, CA) that was set to 0 delay and 10 sec integrate. Fifty µl of Luciferase Assay Reagent II was added to the sample, and the firefly luciferase activity was determined for the pGL2-Promoter (Table 2.2) to normalize for DNA uptake. The *Renilla* luciferase activity was determined after the addition of 50 µl Stop and Glo Reagent.

2.14 IMMUNOBLOT ANALYSIS

Transfected cells were harvested in radioimmunoprecipitation assay (RIPA) buffer (50 mM Tris-HCl [pH 7.4], 150 mM NaCl, 1% Triton-X 100, 1% sodium deoxycholate, 0.1% SDS, 1 mM phenylmethylsulfonyl fluoride [PMSF], 2 mM DTT, 100 µg/ml Leupeptin, 1 µg/ml Pepstatin A, and 40 µg/ml Aprotinin). Debris was removed by centrifugation at 13000 rpm for 5 min at 4°C, and the protein concentration was determined using the Bio-Rad Protein assay. Fifty µg of sample was loaded onto an 8% to 10% Polyacrylamide Gel Electrophoresis (PAGE) gel containing 0.1% SDS and transferred onto a nitrocellulose membrane. The membrane was then blocked with 5% milk in Tris-Buffered Saline Tween-20 (TBST) (20 mM Tris-HCl [pH 7.6], 137 mM NaCl, and 0.05% Tween-20) for 1 hr at 25°C with shaking. The primary antibody was diluted in TBST containing 1% milk and incubated with the membrane for 1 hr at 25°C with shaking. The membrane was washed 3 times for 5 min each in TBST with shaking, and the horseradish peroxidase antibody was then diluted in TBST containing 1% milk and incubated with the membrane for 45 min at 25°C with shaking. The membrane was washed again 3 times for 5 min each with TBST. Western Lightning enhanced chemiluminescent reagent (Perkin Elmer, Wellesley, MA) was used to detect antibody binding.

2.15 SITE-DIRECTED MUTAGENESIS

Mutation of a specific site in a large plasmid was performed using a modified protocol of the Stratagene QuickChange Site-Directed Mutagenesis kit adapted by Wang and Wilkinson (241). A pair of complementary oligonucleotides with the sense primer (5'- GTG CTT ATG AT**A** TTC CC**G** ATT GT**G** TT**T** CA**A** TGC CTT GCG AAG AGC C-3') was synthesized to contain the mutations (in bold) with at least 10 bases of homology to the template plasmid on each end of the primers. The PCR mixture contained 125 ng each oligonucleotide, 50 ng template, 5% DMSO, and 2.5 U *PfuTurbo* DNA polymerase (Stratagene). The PCR conditions were: denaturation at 95^{o}C for 30 sec, and 18 cycles of denaturation at 95^{o}C for 30 sec, annealing at 60^{o}C for 1 min, and extension at 68^{o}C for 14 min in a PTC-100 programmable thermal controller. The DNA was digested with 40 U *Dpn*I for 4 hours at 37^{o}C to remove the original plasmid, and ten µl of the PCR was used to transform *E. coli* DH5α cells.

3. Results

3.1 DEVELOPMENT OF A BACTERIAL DELIVERY SYSTEM

3.1.1 *Shigella* infections

Gene therapy requires vector delivery to cells for subsequent expression of a therapeutic protein (170,209,249). One potential delivery method utilizes vector-containing bacteria to target and invade cells and, once inside, bacterial lysis releases the bacterial cell contents, including the vector, into the eukaryotic cell. Previous studies have shown that an *asd* mutant of *Shigella*, which requires the prokaryotic amino acid diaminopimelic acid (DAP), invades eukaryotic cells but lyses and delivers its contents within the cytoplasm due to the absence of DAP (207). Delivery of a green fluorescent protein (GFP) to human epithelial cells using the *asd* mutant was previously demonstrated by S. Seliger (personal communication). However, to use the MMTV trafficking ability to the mammary gland, the vector must be released into B and T cells. Therefore, to determine if the *Shigella asd* mutant could deliver a vector to mouse or human B and T cells, a modified infection protocol was used for suspension cells.

Four B and T cell lines were assayed for susceptibility to *Shigella asd* infection, and GFP constructs with various promoters were tested for delivery (Table 3.1). None of the cell lines examined showed significant GFP expression

with any of the GFP vectors (data not shown). To verify functionality of the promoter, a *Shigella asd* mutant transformed with a GFP vector containing a T-cell specific promoter from Type B leukemogenic virus (TBLV) was used to infect Jurkat T cells, and the percentages of cells expressing GFP were compared to results from a transient transfection using the same vector constructs. GFP expression was observed at both 24 and 48 hr following infection, but relatively little expression (~0.1%) was observed at either time point (Fig. 3.1). By comparison, the transient transfection showed about 5% expression in cells for both time points, a 50-fold increase in GFP expression relative to the infection. These results suggested low expression was due to the *Shigella* delivery method rather than the promoter.

Table 3.1. Plasmids utilized in bacterial infections.

Plasmid	Characteristics
pJZ466	Eukaryotic GFP expression vector containing an IRES-GFP
pCD3GFP	Eukaryotic GFP expression vector carrying the SV40 promoter
pCMV-EGFP	Eukaryotic GFP expression vector carrying the CMV promoter
p19TLTR-EGFP	Eukaryotic GFP expression vector carrying the TBLV LTR as the promoter
pSS351	Eukaryotic GFP expression vector carrying the CMV promoter and *lacO* sites
pSS356	*yopE* promoter region and first 49 codons fused to *lacI*
pLR56	Prokaryotic vector that expresses invasin
pHYB-MTV	MMTV hybrid provirus in a pBR322 derivative
pIB29MEK	*Y. pseudotuberculosis* virulence plasmid containing *yopH, yopM, yopE,* and *yopK*
pLR71	Prokaryotic GFP expression vector

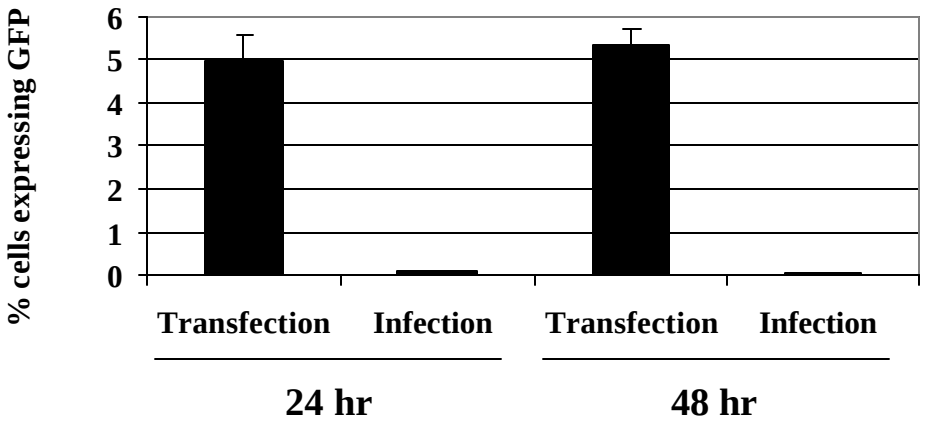


Fig. 3.1. Expression of GFP in Jurkat T cells.

Jurkat T cells were either transiently transfected with p19TLTR or infected with a *Shigella asd* mutant containing p19TLTR. Cells were harvested at 24 or 48 hr, and the percentages of cells expressing GFP are shown as means of triplicate transfections or infections with standard deviations.

Next we determined if there was a problem with vector delivery after bacterial infection. To examine this possibility, transient transfections and bacterial infections both were performed on Jurkat T cells, and the cells were fractionated into cytoplasmic and nuclear DNA fractions. PCR was performed on 5-fold dilutions of each fraction, and the relative levels of DNA from the fractions were compared (Fig. 3.2). A 25-fold increase in delivery efficiency was observed between the transient transfections and the bacterial infections for both the cytoplasmic and nuclear fractions. These results revealed the inefficiency of the *Shigella asd* mutant as a delivery system and prompted development of an alternative system.

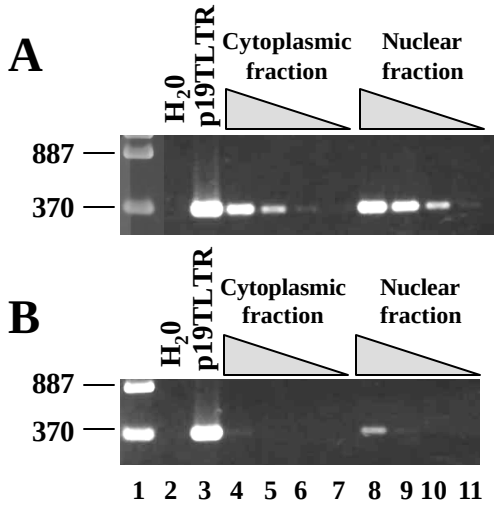


Fig. 3.2. Semi-quantitative PCR analysis of vector localization in the cytoplasmic and nuclear fractions.

(A) DNA from transiently transfected Jurkat T cells was subjected to 5-fold serial dilutions beginning with 5 ml of a 1:125 dilution in lanes 4 and 8. (B) DNA from Jurkat T cells infected with the *Shigella asd* mutant transformed by p19TLTR was subjected to 5-fold serial dilutions. Each set of dilutions was used for PCRs with GFP primers (EGFP for/EGFP rev). All PCRs were analyzed by agarose gel electrophoresis.

3.1.2 *Yersinia* infections

The type III secretion system (TTSS) of *Yersinia* allows secretion and translocation of proteins into the cytoplasm of eukaryotic cells through a pore created by the bacteria (41,42,124). To utilize the TTSS as a gene therapy delivery method, a modified translocation system was constructed and used to transport a GFP vector into eukaryotic cells for expression. *Yersinia* containing the pSS356 plasmid (Table 3.1) should produce a protein with a truncated YopE protein containing the secretion and translocation signals fused to the *lac* repressor (*lacI*) (Fig. 3.3). The GFP vector used in the *Yersinia* experiments (pCMV-EGFP, Table 3.1) was constructed to contain *lac* operators that are bound by *lac* repressor, allowing this plasmid to be transported into the eukaryotic cell for expression during translocation of the YopE-LacI protein. Other studies from Dr. Shelley Payne's lab have shown delivery of GFP to human epithelial cells using this method (personal communication).

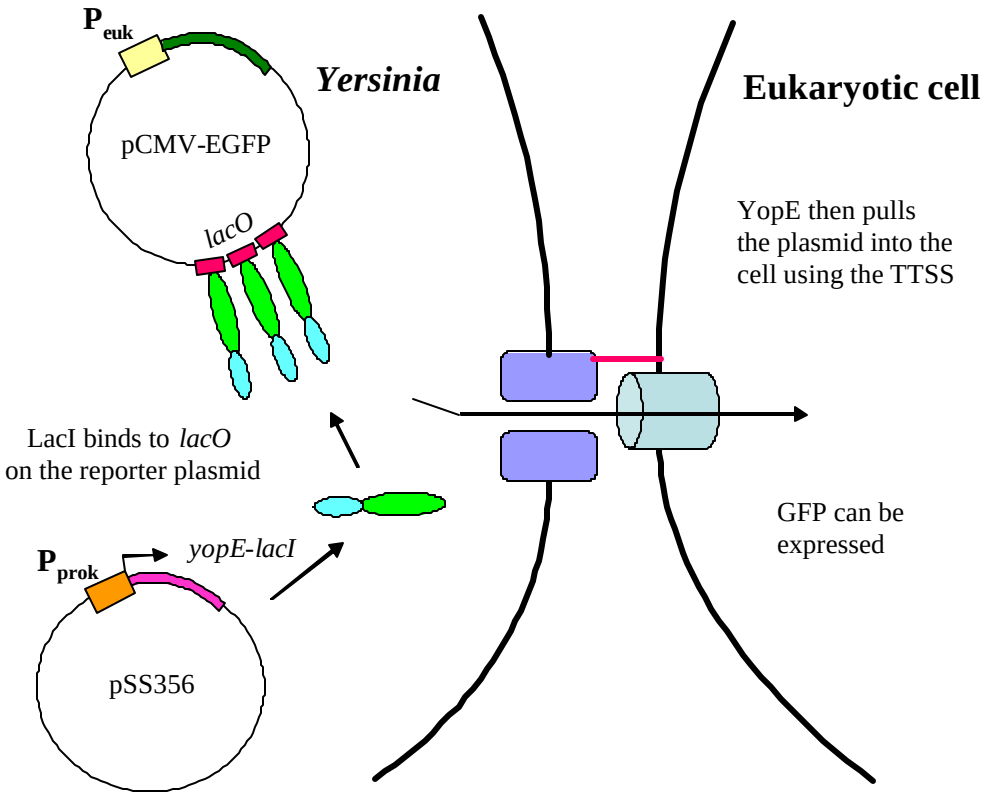


Fig. 3.3. Strategy for delivery of GFP plasmids to eukaryotic cells using the type III secretion system of *Yersinia*.

The *Yersinia* contain a plasmid with a prokaryotic promoter that allows expression of a YopE-LacI fusion protein. The LacI can bind to *lacO* sites located on GFP plasmids also present in the *Yersinia*. The YopE uses its secretion and translocation signals to move into the cytoplasm of the eukaryotic cell through a pore of the type III translocation system. The GFP plasmid also should move into the cytoplasm due to LacI binding to the *lacO*.

To determine if this system could deliver a GFP vector to lymphocytes so that the MMTV trafficking ability to the mammary gland could be utilized, modified infections were performed on a panel of three mouse and human B and T cell lines. None of the cell lines showed detectable GFP expression. Unexpectedly, when a linear scale was used to set the analysis gate for the lymphocyte population, two different populations were observed (Fig. 3.4). The circled population contained cells that were the normal population to be gated for analysis, while the other population showed cells with a different size and granularity than those of the normal population. This other population most likely consists of dead or dying cells; however, these cells also might consist of living cells that are undergoing morphological changes due to the addition of the bacteria. This other cell population for these cell lines was unusually high compared to other bacterial infection analyses, and these data suggested that the *Yersinia* infection caused toxicity in a variety of cell types.

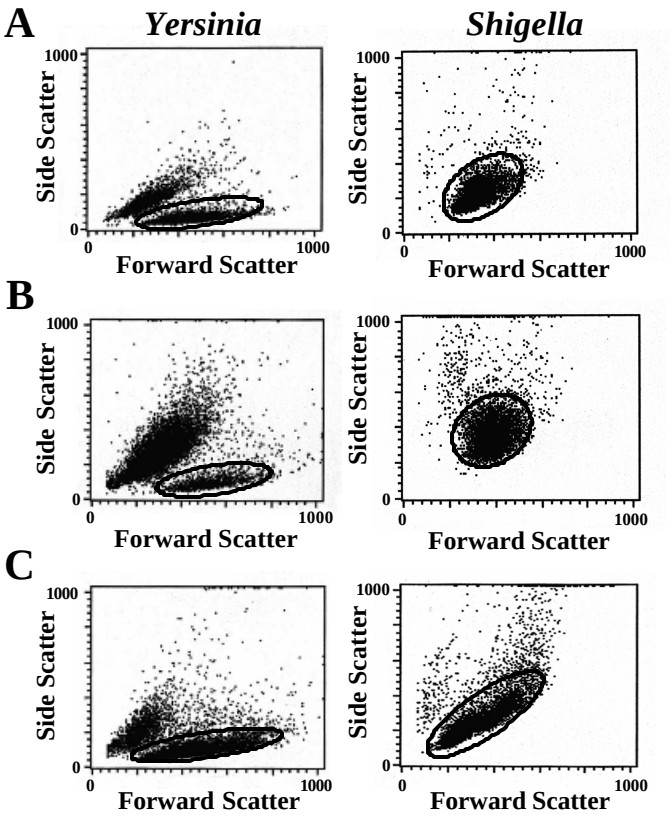


Fig. 3.4. Analysis of the population of B and T cells after bacterial infection.

Linear forward and side scatter plots are shown for (A) A20 B cells, (B) LBB.11 B cells, and (C) RLmale1 T cells. For each, the first panel shows an infection by *Yersinia*, and the second panel shows an infection by *Shigella* harvested at 48 hr. The viable cell population is indicated with a black circle.

Since Dr. Payne's lab had previously obtained GFP analysis results from *Yersinia* infection of HeLa cells and did not report cell toxicity, we repeated their analysis conditions. The infection was assayed at 48 hr, and forward and side scatter FACS analysis was performed on both linear and log scales. Most cell types are routinely analyzed on a linear scale as was previously performed on the HeLa cell infection. After infection, the HeLa cells were lifting off the plate and, therefore, insufficient attached cells were available to harvest and analyze for the correct population or GFP expression; however, the unattached cells were still harvested for analysis. The unattached HeLa cells appear to be a normal population that contains living cells (circled population) (Fig. 3.5A). However, for the same unattached HeLa cells, the logarithmic plot distinguishes two distinct populations (Fig. 3.5B). The circled population contains the live cells that should be analyzed, while the majority of the cells are located in another population consisting of dead cells and debris.

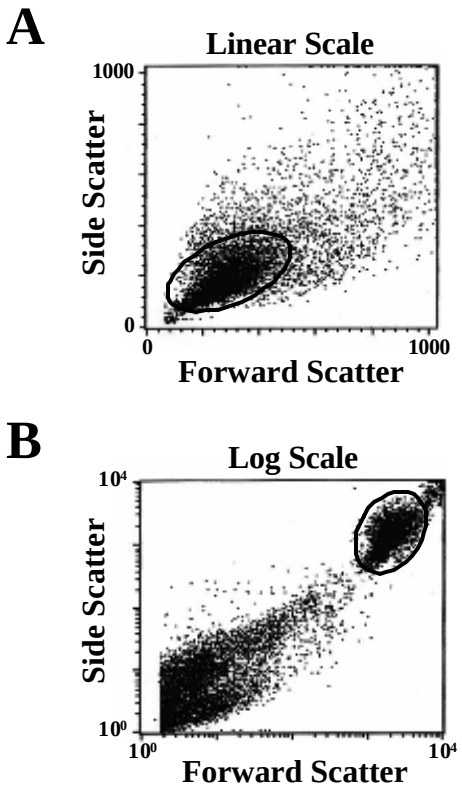


Fig. 3.5. Analysis of the population of HeLa cells after infection by *Yersinia*.

(A) HeLa cells are shown using a linear forward and side scatter plot. (B) HeLa cells are shown using a logarithmic forward and side scatter plot. The normal population of cells is indicated with a black circle.

The experiment also was modified to use shorter bacterial incubation times that potentially would reduce cell death. The HeLa cells were harvested at both 24 and 48 hr to determine if and when cell death was occurring, and the mock-infected cells were analyzed with and without cytochalasin D (CCD), which prevents bacterial uptake into the cell. At both time points, the mock cell populations looked similar and appeared to be a normal population, suggesting that CCD does not affect the viability of HeLa cells (Fig. 3.6). Twenty-four hr after *Yersinia* infection, the cell populations were starting to disperse, but a population could still be observed and analyzed (circled population). However, 48 hr after the addition of the *Yersinia*, a tight population is no longer seen, and the cells are quite dispersed. These data and previous results suggested that *Yersinia* infection of eukaryotic cells causes cytotoxicity.

In addition to observing cell death, GFP expression also was analyzed. Little expression was observed, and the results were inconsistent when examined by gating on the normal population seen in the mock-infected cells (black circle) (Fig. 3.6). This outcome was mostly likely due to the unusual population of cells after the addition of the bacteria. GFP expression also was observed on the entire ungated population to determine the effect of autofluorescence from the dead or dying cells. After 24 hr, when the cells still appeared as a normal population, no GFP expression was seen (Fig. 3.7). These results suggested that the cells were

dying, and no delivery of the vector had taken place. After 48 hr, when the cells were dispersed and later in the dying process, a peak separate from the mock peak appeared, which normally would indicate delivery of the GFP vector. This peak also was seen in earlier studies, yet these data were not supported by the gated results. Analysis of gated cells suggested that the peak was due to autofluorescence from the dead cells. Since we were unable to prevent the cell lysis seen during the *Yersinia* infections, a third gene delivery system was tested.

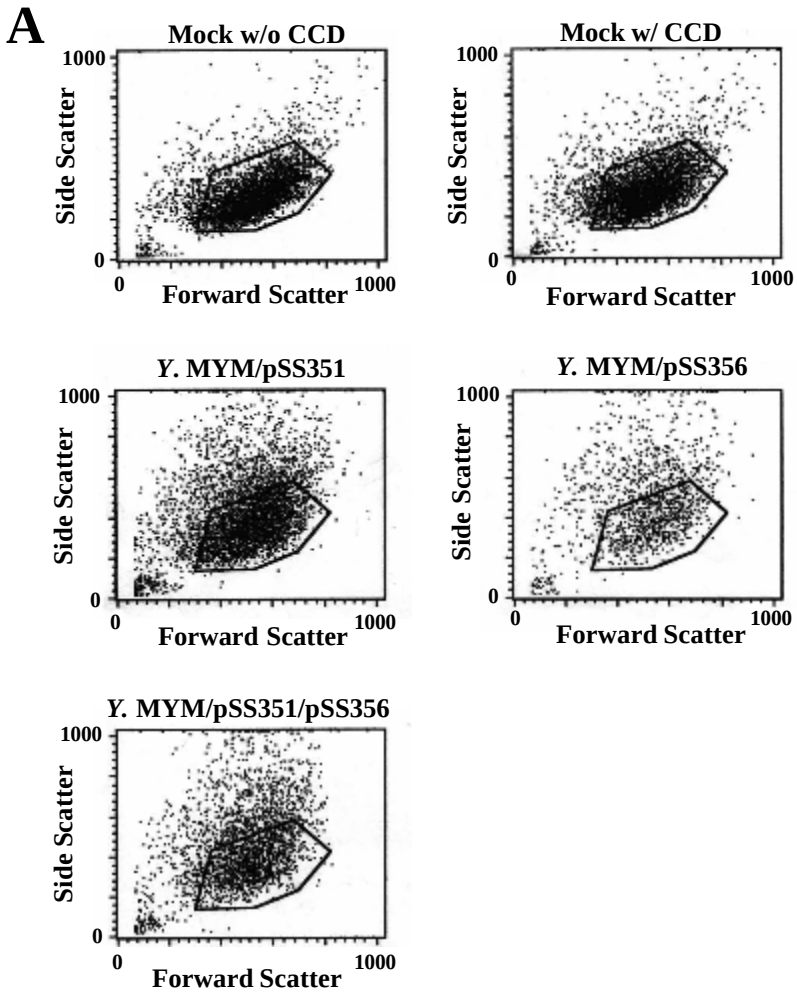
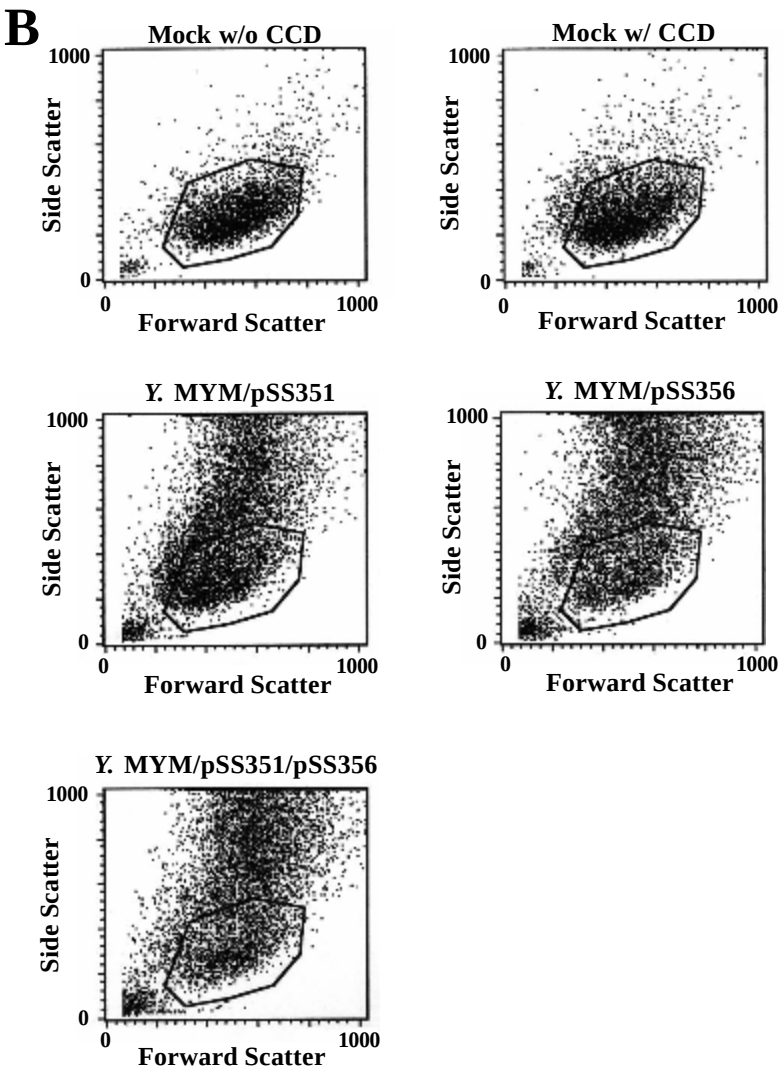


Fig. 3.6. Analysis of the population of HeLa cells after infection by *Yersinia* at different time points.

HeLa cells are shown using a linear forward and side scatter plot after (A) 24 hr and (B) 48 hr. The *Yersinia* used to infect the HeLa cells are indicated above the panels. Mock-infected HeLa cells with and without CCD were used as a negative control for the cells and the drug. The normal population of cells is indicated with a black circle. Plasmid pSS351 is the eukaryotic expression vector carrying CMV-GFP and *lac*O sites, whereas pSS356 is the prokaryotic vector carrying the *yopE* promoter upstream of the YopE-LacI fusion.


B
Mock w/o CCD
Mock w/ CCD
Y. MYM/pSS351
Y. MYM/pSS356
Y. MYM/pSS351/pSS356
Side Scatter
Forward Scatter
1000
0

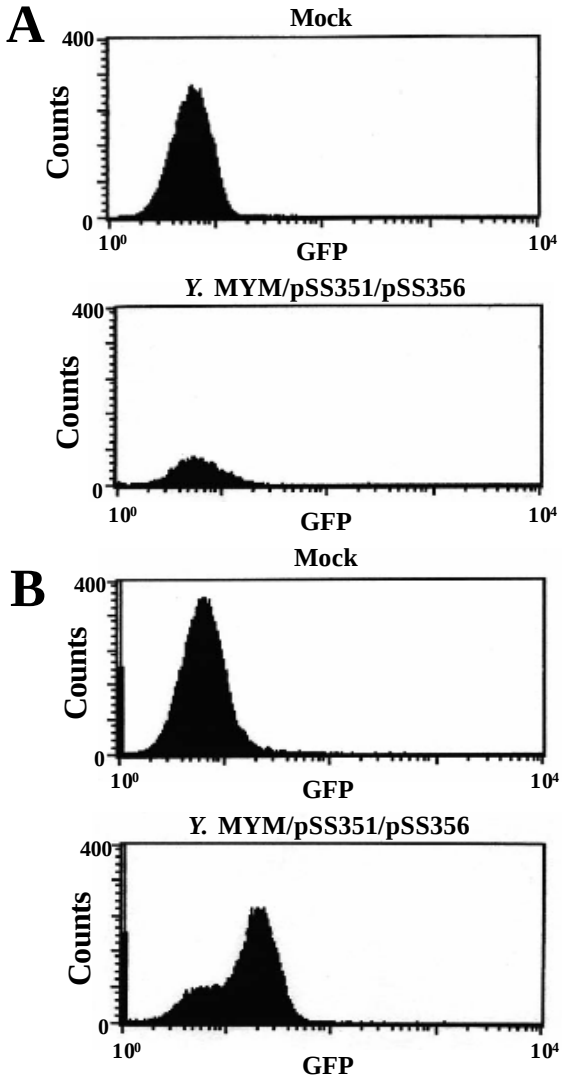


Fig. 3.7. Expression of GFP in HeLa cells after infection by *Yersinia* at different time points.

HeLa cells are shown using a histogram plot after (A) 24 hr or (B) 48 hr. The *Yersinia* used for HeLa infection are indicated above the panels. GFP expression is shown as a shift to the right of the peak of mock-infected cells.

3.1.3 *E. coli* infections

The genetics of *E. coli* are widely known and can easily be manipulated (20,35). Therefore, this system was used in combination with the TTSS of *Yersinia* to deliver vectors to eukaryotic cells. To determine whether *E. coli* would kill eukaryotic cells, a preliminary infection of HeLa epithelial cells and A20 B cells was performed. *E. coli* transformed with different plasmids as well as an *invA* plasmid that produces invasin needed for attachment to β1 integrins on cells (described in Table 3.1) were used for infection. Both the HeLa cells and the A20 B cells infected with *E. coli* showed no observable increase in cell death when compared to the mock-infected cells, with the exception of the A20 B cells infected with the *E. coli* containing pHYB-MTV, an infectious MMTV provirus (Fig. 3.8). However, the 10% cell death observed was 5-fold lower than that of the A20 B cells infected with *Yersinia*. In addition, the HeLa cells were analyzed with propidium iodide, which is a fluorescent dye that is only incorporated into the DNA of dead or dying cells (215). At 48 hr post infection, the amount of fluorescent dead cells was similar (~30%) between the mock-infected cells and the bacterially-infected cells (Fig. 3.9). Together, these results suggest that *E. coli* does not kill eukaryotic cells within 48 hr.

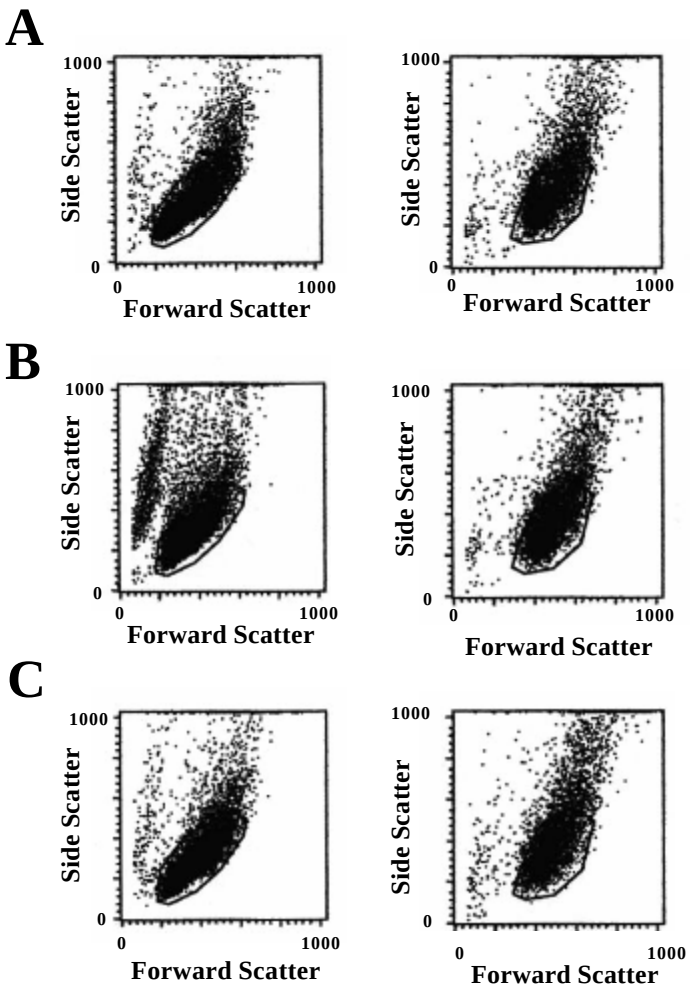


Fig. 3.8. Analysis of the population of A20 cells and HeLa cells after infection by *E. coli.*

A forward and side scatter plot is shown for (A) mock-infected cells, (B) DH5α/pHYB-MTV-infected cells, and (C) DH5α/pIB29MEK-infected cells. The first panel in each set shows A20 B cells, and the second panel shows HeLa epithelial cells. The mock-infected cells were used as a negative control, and the normal population of cells is indicated with a black circle. pHYB-MTV contains an infectious MMTV provirus, whereas pIB29MEK is a *Yersinia* virulence plasmid containing mutations in *yopH*, *M*, *E*, and *K*.

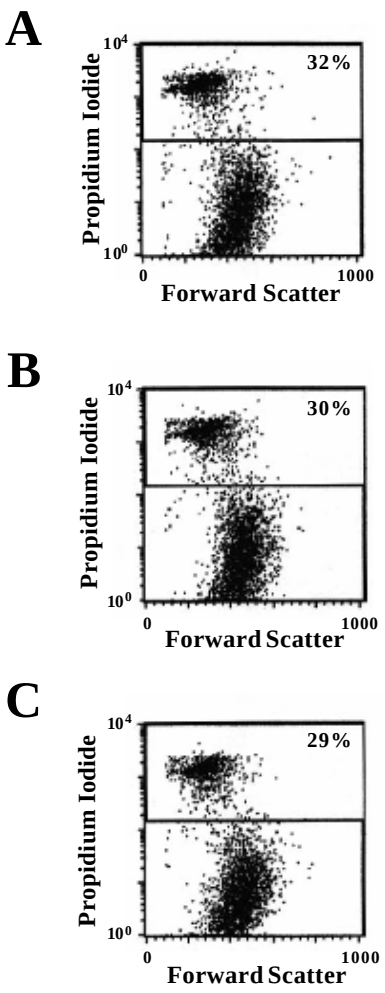


Fig. 3.9. Propidium iodide analysis of HeLa cells after infection by *E. coli*.

HeLa cells were either (A) mock-infected, (B) infected with DH5α/pHYB-MTV, or (C) infected with DH5α/pIB29MEK and harvested at 48 hr. Each panel shows the infected HeLa cells treated with propidium iodide with the normal population of cells located in the lower rectangle. Propidium iodide incorporation into dead cells is shown as a shift upward, and the percentages of these cells are indicated in the top right corner of the panel. pHYB-MTV contains an infectious MMTV provirus, whereas pIB29MEK is a *Yersinia* virulence plasmid containing mutations in *yopH, M, E,* and *K*.

Although *E. coli* did not kill the cells, it was unclear whether the bacteria attached to the cells using the invasin provided from *Yersinia*. To determine if attachment occurred, the *invA* gene was integrated into the *E. coli* chromosome (DH5α/*inva*), and the bacteria were transformed with a prokaryotic GFP expression plasmid (pLR71, Table 3.1). The DH5α/*inva* bacteria then were utilized for infections of HeLa epithelial cells and lymphocytes, and infections with *E. coli* carrying the GFP plasmid, but without invasin, were used as controls (DH5α/pLR71). The cell lines used for infection expressed β1 integrins, to which the invasin could bind, on their surface (data not shown), so it was expected that cells would show an increase in fluorescence when the bacteria contained the invasin gene. Surprisingly, with both the HeLa cells and the RLmale1 T cells, the same change in fluorescence was observed even after infection with the *E. coli* that did not contain the invasin gene (Fig. 3.10). Despite this, approximately 10% of HeLa cells had attached bacteria, while 100% of RLmale1 cells showed bacterial attachment. These results suggested that either the *E. coli* were able to attach to the cells by a means other than the invasin, or that non-specific binding between the bacteria and the cells had occurred.

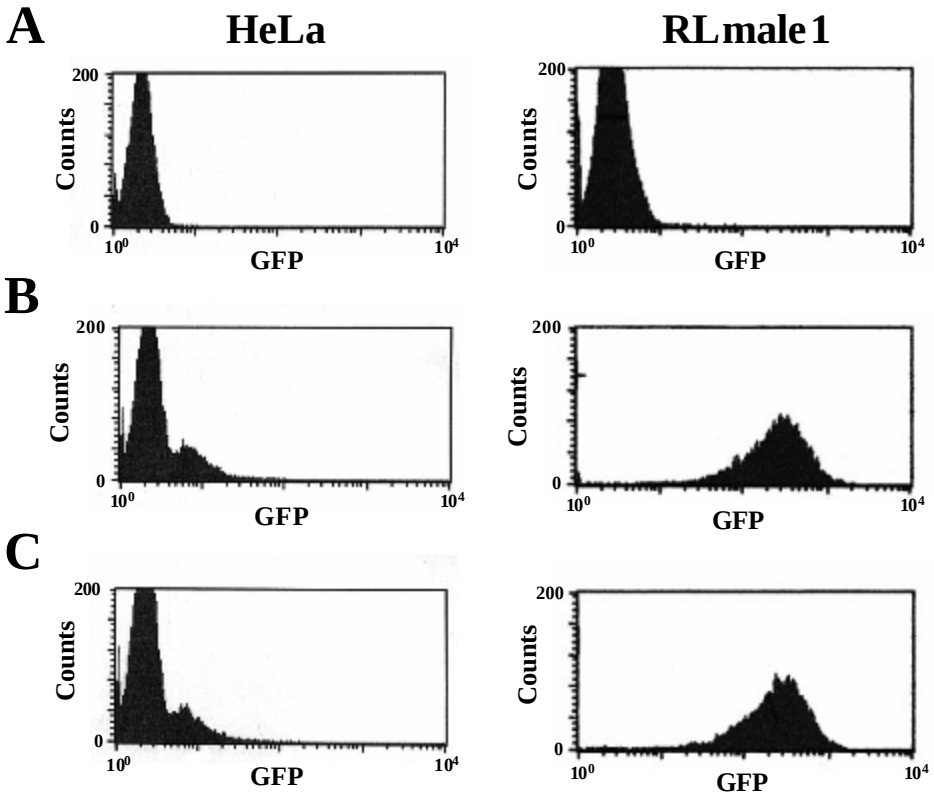


Fig. 3.10. Attachment of fluorescent bacteria to human HeLa or mouse RLmale1 cells after infection by *E.coli*.

A histogram plot is shown for (A) mock-infected cells, (B) DH5α/pLR71-infected cells, and (C) DH5α/*inva*/pLR71-infected cells harvested at 48 hr. The first panel in each set shows HeLa epithelial cells, and the second panel shows RLmale1 T cells. The mock-infected and DH5α/pLR71-infected cells were used as negative controls. Bacterial attachment to the eukaryotic cells is shown as a shift to the right of the peak of mock-infected cells.

To determine the level of nonspecific binding, the bacteria and cells were centrifuged through a Histopaque gradient that pelleted the bacteria and allowed the eukaryotic cells to remain at the interphase. For the RLmale1 T cells, the infection without the Histopaque was the same as reported earlier (compare Fig. 3.10 and 3.11A). When the infection was repeated using the Histopaque washing, some of the non-specific binding was detected, but the majority of the cells appeared to bind *E. coli* irrespective of the invasin (Fig. 3.11B). To determine if another protein or mechanism was being used for attachment, the EL4b T cell line, which does not express β1 integrins, was used for the infection. The results observed for EL4b cells without Histopaque washing were comparable to those observed for the RLmale1 cells, i.e., all the eukaryotic cells had attached bacteria (Fig. 3.12A). To remove unattached bacteria, the infection was repeated with Histopaque washing. Surprisingly, although 50% of the cells had non-specifically bound bacteria, half of the eukaryotic cells showed bacterial attachment that could not be attributed to the invasin (Fig. 3.12B). These results demonstrate that *E. coli* attach to eukaryotic cells through at least two mechanisms. Further modifications of this system will be required to allow specific targeting of eukaryotic cells.

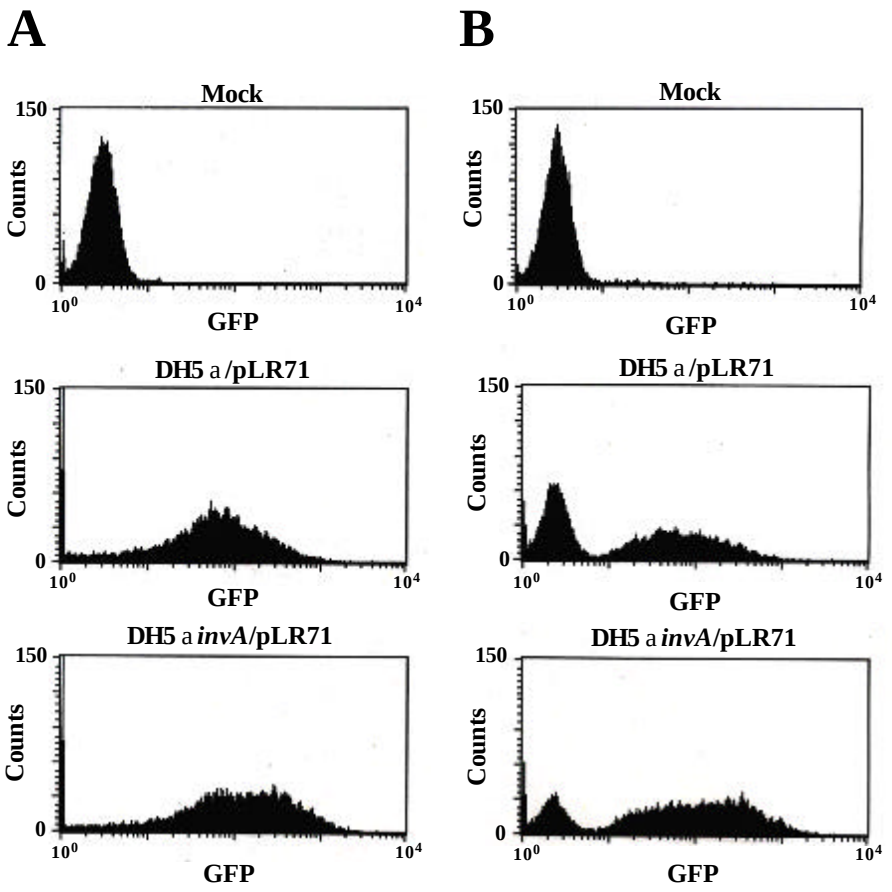


Fig. 3.11. Attachment of fluorescent bacteria to RLmale1 mouse T cells after infection by *E. coli* measured by different washing techniques.

A histogram plot is shown for the infected RLmale1 cells washed with (A) phosphate-buffered saline or (B) Histopaque and harvested at 48 hr. The *E. coli* used to infect the RLmale1 cells are indicated above the panels. The mock-infected and DH5α/pLR71-infected cells were used as negative controls. Fluorescent bacteria attached to the cells are shown as a shift to the right of the peak of mock-infected cells.

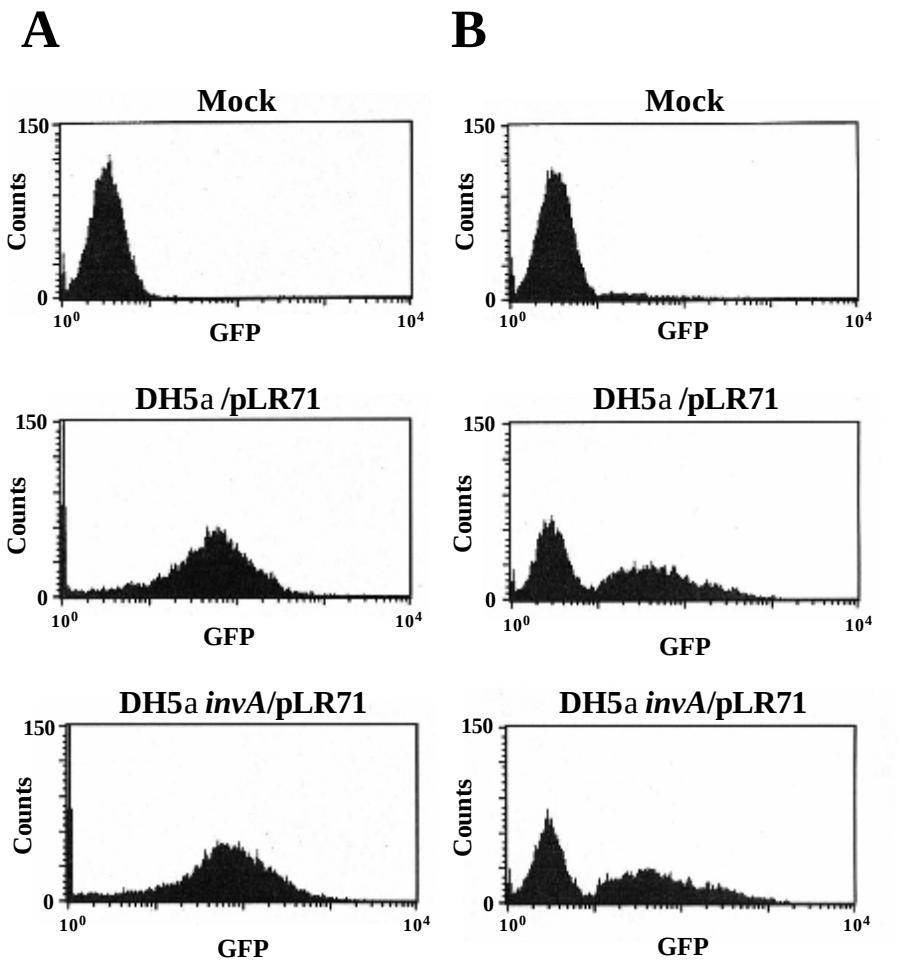


Fig. 3.12. Attachment of fluorescent bacteria to EL4b mouse T cells after infection by *E. coli* measured by different washing techniques.

A histogram plot is shown for the infected EL4b cells washed with (A) phosphate-buffered saline or (B) Histopaque and harvested after 48 hr. The *E. coli* used to infect the EL4b cells are indicated above the panels. The mock-infected and DH5α/pLR71-infected cells were used as negative controls. Fluorescent bacteria attached to the cells are shown as a shift to the right of the peak of mock-infected cells.

3.2 CONSTRUCTION AND CHARACTERIZATION OF A GFP-TAGGED MMTV

3.2.1 Analysis of GFP expression from the tagged MMTVs

MMTV is a retrovirus that causes mammary gland carcinomas following infection and amplification in B and T cells (6,26,56,70). Since MMTV first infects these immune cells, lymphocyte trafficking or homing must occur before infection of the mammary gland. Although much is known about lymphocyte homing to specific tissues (117,230,247), the mechanism of MMTV transmission from gut-associated lymphocytes to the mammary gland is not well understood. Use of MMTV as a gene therapy vector requires additional details of virus-cell interactions.

To examine MMTV trafficking from infected lymphocytes to the mammary gland, a GFP-tagged MMTV was constructed as described in the Materials and Methods. Briefly, an IRES-GFP was engineered 3' to the highly expressed *env* gene after duplication of the sequence shared by both *env* and *sag* (refer to Materials and Methods). Duplication of these shared sequences would allow GFP translation from the *env* mRNA without truncating the essential *sag* mRNA. The GFP-tagged MMTV clones initially were tested for GFP expression by transient transfection of XC fibroblast cells (Fig. 3.13). The positive control plasmid containing the IRES-GFP used for cloning (pJZ466, Table 3.2) showed

that approximately 17% of the XC cells expressed GFP. Transfection of two GFP-tagged MMTV plasmid isolates, pHYB-GFP.1 and pHYB-GFP.8, showed approximately 3% and 4.5% of XC cells expressing GFP, respectively. These results suggested that the GFP-tagged MMTV clones expressed GFP, but at 3- to 5-fold lower levels than the GFP expression plasmid used as a positive control.

Table 3.2. Plasmids utilized in the construction and characterization of GFP-tagged MMTV.

Plasmid	Characteristics
pJZ466	Eukaryotic GFP expression vector containing an IRES-GFP
pHYB-GFP.1	One clone of the pHYB-MTV derivative containing an IRES-GFP 3' to the *env* gene
pHYB-GFP.8	Another clone of the pHYB-MTV derivative containing an IRES-GFP 3' to the *env* gene
pMMTV-GFP	Eukaryotic GFP expression vector carrying the C3H MMTV LTR as the promoter
pHYB-GFPmSA	pHYB-GFP.1 derivative containing a mutated splice acceptor preceding the IRES-GFP
pHYB-GFPmSA.bs	pHYB-GFPmSA derivative containing a branch site and polypyrimidine tract preceding the last splice acceptor site

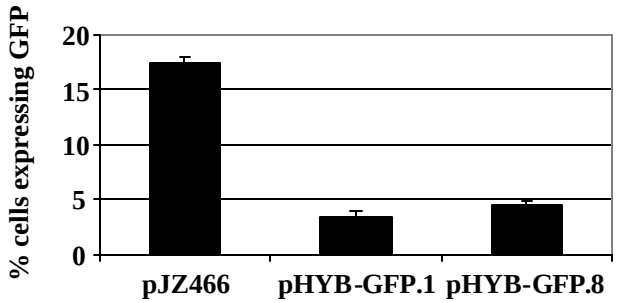


Fig. 3.13. Expression of GFP in XC cells from the tagged-MMTVs.

XC fibroblast cells were transiently transfected with either pJZ466 or GFP-tagged MMTVs. Cells were harvested at 48 hr, and the percentages of cells expressing GFP are shown as means of triplicate transfections with standard deviations.

The MMTV *env* mRNA is transcribed from the hormone-inducible U3 promoter (146,227). Since GFP also is expressed from the *env* mRNA, we determined if GFP expression could be induced by the glucocorticoid, dexamethasone (Dex). XC cells were transiently transfected, and 10^{-6} M Dex was added to the cells for 24 hr prior to fluorescence assays. The plasmid pMMTV-GFP (Table 3.2), which uses the MMTV promoter, was used as a positive control and had a 3-fold increase in GFP expression in the presence of Dex (Fig. 3.14). The clones pHYB-GFP.1 and pHYB-GFP.8 also had a 3-fold and 6.5-fold increase in GFP expression, respectively, upon addition of Dex, whereas a non-inducible plasmid, pJZ466 (Table 3.2), showed no increase in GFP expression after Dex induction (data not shown). These results demonstrated that GFP expression is Dex inducible and most likely originates from the hormone-inducible U3 promoter.

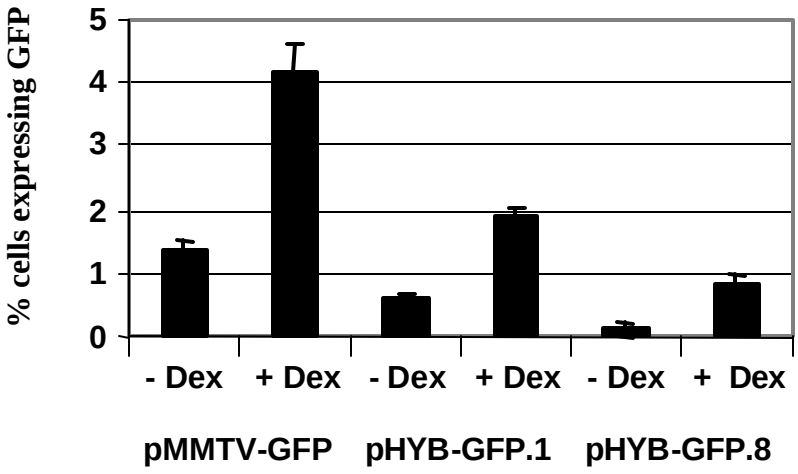


Fig. 3.14. Expression of GFP from the tagged-MMTVs after Dex induction in XC cells.

XC fibroblast cells were transiently transfected with either pMMTV-GFP, which expresses GFP under the control of the MMTV LTR, or GFP-tagged MMTV mutants. Dex (10^{-6} M) was added to the cells after 24 hr, and the cells were harvested at 48 hr. The percentages of cells expressing GFP are shown as means of triplicate transfections with standard deviations.

3.2.2 Expression of viral transcripts from GFP-tagged MMTVs

Since both *env* and *sag* are required for a productive MMTV infection, functional proteins also will be required for use of MMTV as a gene therapy vector. To verify that the duplication of the *env* and *sag* sequence as well as the insertion of an IRES-GFP had not disrupted expression of either of these genes, RT-PCR was performed on total RNA harvested from transiently transfected XC cells to detect viral mRNAs. Primers designed to detect *gag*, *env*, and *sag* mRNAs, as well as primers to detect the IRES within the *env* mRNA, were used in the RT-PCR analysis (Fig. 3.15). The positive control pHYB-MTV as well as the GFP-tagged MMTV clones all produce *gag* and *env* mRNAs (Fig. 3.16). Surprisingly, only pHYB-MTV showed a product for the *sag* mRNA, suggesting that the pHYB-GFP clones do not make a *sag* transcript. Since wild type MMTV does not contain an IRES-GFP, pHYB-MTV does not show a band for the *env-IRES* mRNA. The expected amplification product for the *env-IRES* mRNA was approximately 2 kb, yet the bands seen with both GFP-tagged MMTV clones are approximately 900 bp. The 900 bp band was isolated and sequenced. Sequence analysis verified that the 900 bp band was a PCR product from the *env* and *IRES* primers used, but suggested that alternative splicing might be occurring.

To further analyze the transcripts detected, the splice donor (SD) and splice acceptor (SA) sites were examined for all the mRNAs. A 2 kb *env-IRES*

transcript from the U3 promoter might be expected after removal of the *gag-pol* intron (Fig. 3.15). The smaller PCR product observed suggested that a doubly-spliced mRNA was being produced (Fig. 3.16D). The *sag* mRNA transcribed from wild-type MMTV is expressed from an intergenic *env* promoter (59) and uses SD 7315 in *env* and a SA site that is located approximately 70 bp before the 3' LTR. During the construction of the GFP-tagged MMTVs, SA 8467 located prior to the IRES-GFP was maintained and an engineered SA was placed between IRES-GFP and the 3' LTR. Since *sag* transcripts were not detected using the GFP-tagged MMTV clones, the RT-PCR results suggest that the original SA is used instead of the engineered SA site.

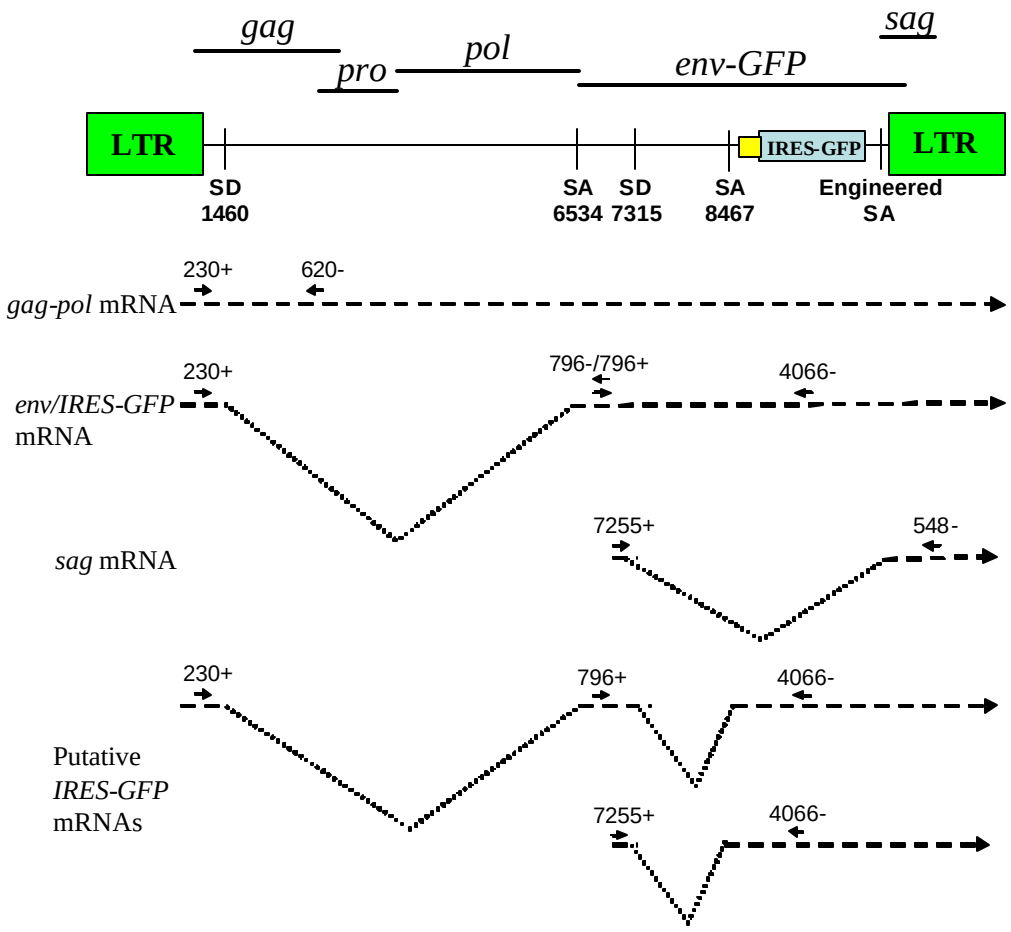


Fig. 3.15. Diagram of the splice donor and splice acceptors in the GFP-tagged MMTV.

The proteins of the GFP-tagged MMTV are shown as horizontal lines above the provirus, and the splice donors and splice acceptors are indicated by vertical lines within the provirus. The viral transcripts are shown by the dotted lines below the provirus, and the primers used for each mRNA are indicated.

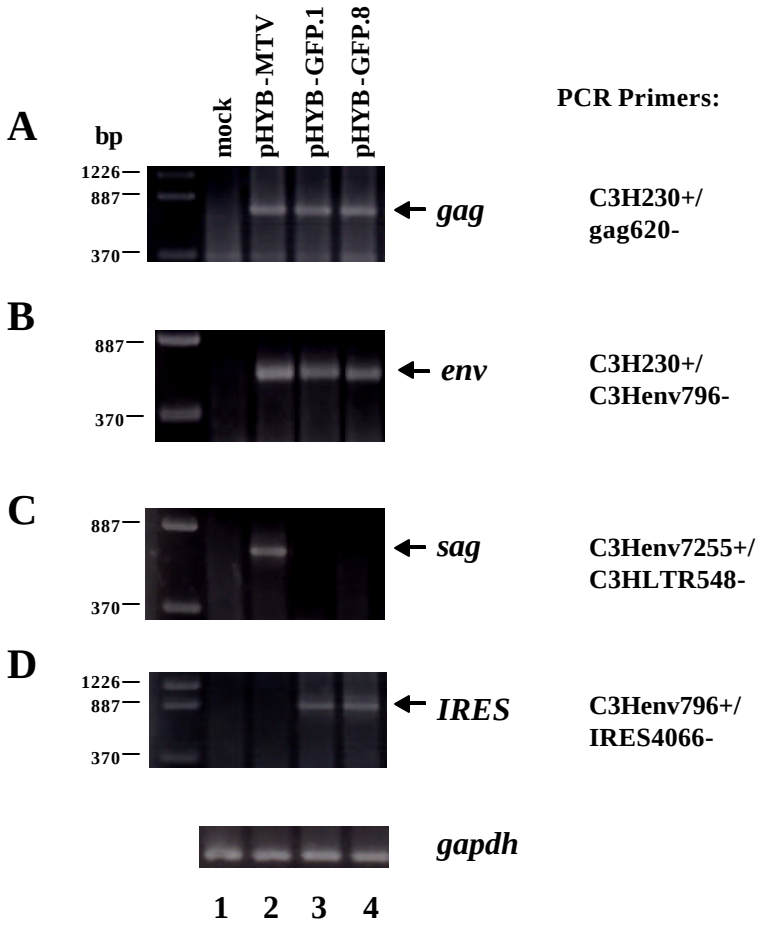


Fig. 3.16. Detection of viral transcripts from XC cells transiently transfected with GFP-tagged MMTVs.

XC fibroblast cells were transiently transfected with either pHYB-MTV or GFP-tagged MMTVs, and total RNA was harvested 48 hr post transfection. cDNA was prepared, and RT-PCR performed for (A) *gag* mRNA, (B) *env* mRNA, (C) *sag* mRNA, and (D) *IRES* mRNA using the indicated primers. Arrows show the locations of the expected products for each primer pair listed. The *gapdh* mRNA was used as a control, and all PCRs were analyzed by agarose gel electrophoresis.

3.2.3 Construction and characterization of a GFP-tagged MMTV with a mutated splice acceptor

The splicing machinery of a eukaryotic cell frequently utilizes the first SA located after a SD site (36). The GFP-tagged MMTV clones also appear to use this common strategy since no *sag* mRNA is being produced. Therefore, to construct a GFP-tagged MMTV with the ability to produce *sag* mRNA, the SA site located before the duplicated sequence (SA 8467) was mutated using a modified protocol for the Stratagene QuickChange Site-Directed Mutagenesis kit. The resulting clone, pHYB-GFPmSA (Table 3.2), was transiently transfected into XC cells to determine if GFP was expressed (Fig. 3.17). The positive control plasmid pJZ466 showed 3.2% GFP expression, while pHYB-GFPmSA only showed 0.2% GFP expression. Although the expression level varied from previous results, probably due to different transfection efficiencies, the GFP expression from pHYB-GFPmSA is considerably lower than that from the previous pHYB-GFP construct. This is not surprising since pHYB-GFP had both a doubly-spliced *IRES-GFP* mRNA and a singly spliced mRNA from the *env* intergenic promoter in addition to the *env-IRES* mRNA from the U3 promoter (Fig. 3.15). However, in pHYB-GFPmSA, two of the mRNAs used to translate GFP were no longer made, thereby decreasing the amount of GFP produced.

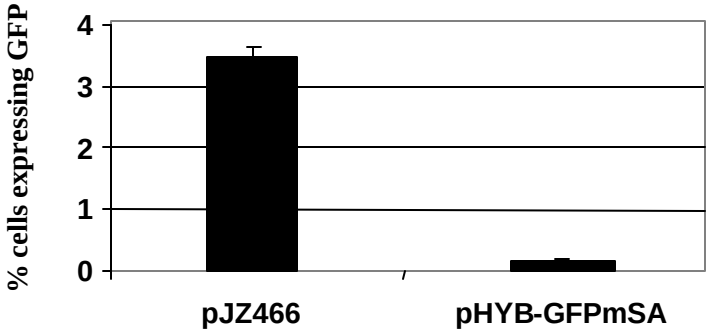


Fig. 3.17. Expression of GFP in XC cells from the tagged-MMTV containing a mutated splice acceptor.

XC fibroblast cells were transiently transfected with either pJZ466 or a GFP-tagged MMTV containing a mutated splice acceptor. Cells were harvested at 48 hr, and the percentages of cells expressing GFP are shown as means of triplicate transfections with standard deviations.

Total RNA also was harvested from transiently transfected XC cells to determine if viral mRNAs could be detected (Fig. 3.18). Similar to results with pHYB-GFP, both *gag* and *env* mRNAs were detectable from the pHYB-GFPmSA transfections. The *env-IRES* mRNA was not detected after mutation of the SA at position 8467. However, the previously detected 900 bp PCR product also was not observed (Fig. 3.18C). The difficulty in detecting the *env-IRES* mRNA was probably due to the PCR conditions and the length of the product. Since the pHYB-GFPmSA construct does not produce a *sag* mRNA, there appear to be additional factors, other than duplication of the SA site, affecting this transcript.

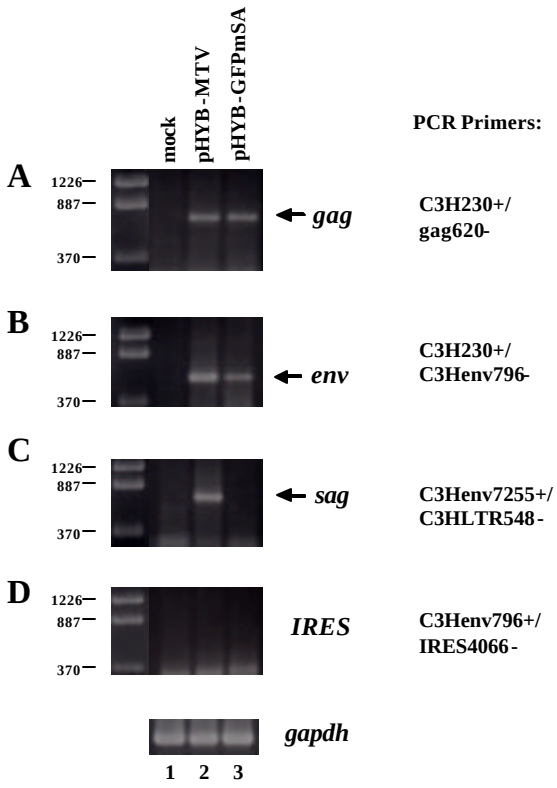


Fig. 3.18. Detection of viral transcripts from XC cells transiently transfected with GFP-tagged MMTV containing a mutated splice acceptor.

XC fibroblast cells were transiently transfected with either pHYB-MTV or pHYB-GFPmSA containing a mutated splice acceptor, and total RNA was harvested 48 hr post transfection. cDNA was prepared, and RT-PCR performed for (A) *gag* mRNA, (B) *env* mRNA, (C) *sag* mRNA, and (D) *IRES* mRNA using the indicated primers. Arrows show the locations of the expected products for each primer pair listed. The *gapdh* mRNA was used as a control, and all PCRs were analyzed by agarose gel electrophoresis.

3.2.4 Construction and characterization of a GFP-tagged MMTV containing splicing regulatory sequences

Previous studies on splicing have shown that sequences other than the SD and SA sites are important for recognition by the splicing machinery of the cell (87,153). These sequences include a polypyrimidine tract as well as a branch site, which is the adenine attacked for splicing to occur (Fig. 3.19). The engineered SA site in pHYB-GFPmSA did not contain any of these additional sequences and may have prevented *sag* mRNA production. Therefore the *sag* mRNA polypyrimidine tract and branch site sequences were inserted before the engineered SA site in pHYB-GFPmSA as follows. Two complementary oligonucleotides were synthesized that contained a branch site, a polypyrimidine tract, a SA site, and flanking restriction enzyme sites. These oligonucleotides were annealed and phosphorylated, and the clone pHYB-GFPmSA.bs was created by substitution of this more complete splice site for that engineered into pHYB-GFPmSA (Table 3.2). To determine if *sag* mRNA was expressed from this construct, XC cells were transiently transfected and total RNA was harvested for RT-PCR analysis (Fig. 3.20). Unexpectedly, pHYB-GFPmSA.bs did not produce a *sag* mRNA, although *env* mRNA was detectable. These results suggested that either the cellular splicing machinery was not functioning as expected, or the mutant MMTV interfered with other *sag* mRNA regulatory sequences.

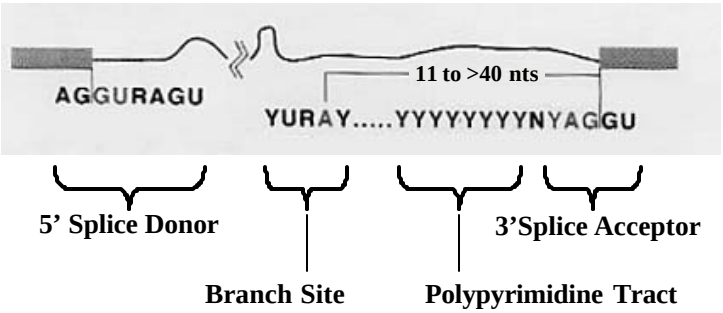


Fig. 3.19. RNA sequences important for recognition by the splicing machinery.

The RNA sequences that are important for recognition by the splicing machinery are shown as consensus sequences from mammalian cells. The 5' splice donor, the branch site, the polypyrimidine tract, and the 3' splice acceptor are all indicated. The branch site is located from 11 to greater than 40 nucleotides before the 3' splice acceptor, and it is the adenine (A) from this branch site that is recognized by the splicing machinery [adapted from Moore , MJ, 2000 (153)].

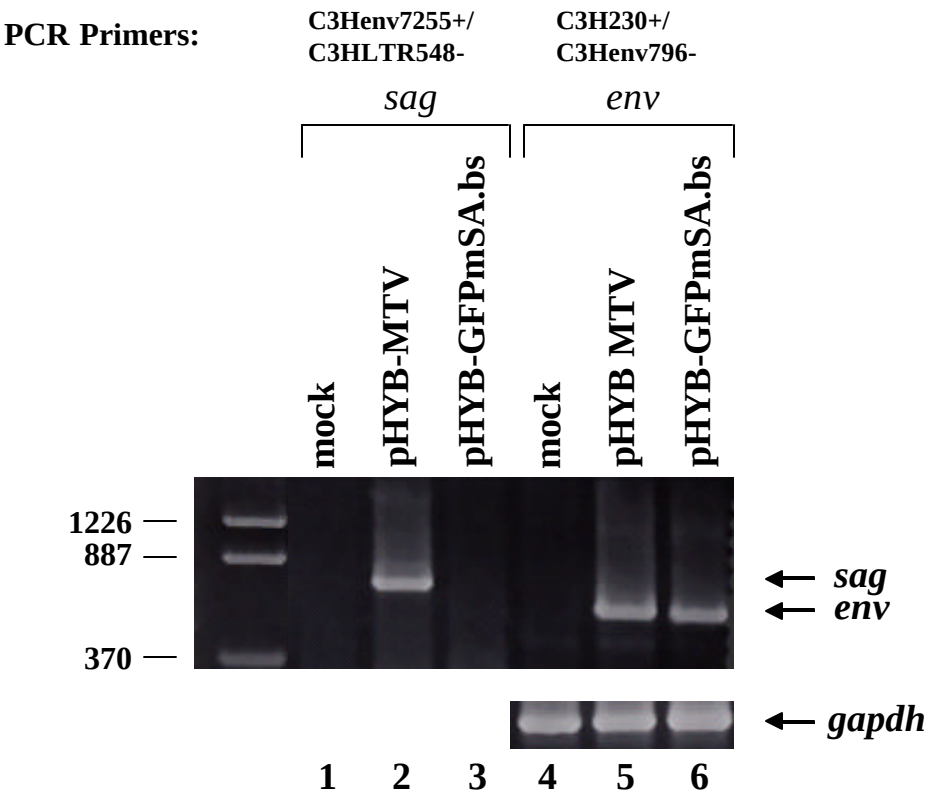


Fig. 3.20. Detection of viral transcripts from XC cells transiently transfected with GFP-tagged MMTV containing splicing regulatory sequences.

XC fibroblast cells were transiently transfected with either pHYB-MTV or GFP-tagged MMTV containing splicing regulatory sequences, and total RNA was harvested at 48 hr post transfection. cDNA was prepared, and RT-PCR was performed using the indicated primers. Arrows show the locations of the expected products for each primer pair listed. The *gapdh* mRNA was used as a control, and all PCRs were analyzed by agarose gel electrophoresis.

3.2.5 Detection of the capsid protein from the GFP-tagged MMTV mutants

The RT-PCR analysis is necessary to detect *sag* mRNA since Sag antibody (Ab) is not available (27,150,181). To determine if viral structural proteins are expressed from the GFP-tagged MMTV proviruses, immunoblot analysis was performed on lysates harvested from transiently transfected XC cells (Fig. 3.21). Polyclonal or monoclonal MMTV CA-specific Ab was used to detect CA from pHYB-MTV transfected cells as a positive control. Using the polyclonal Ab, the expected CA precursor Pr77 protein was observed (lane 2). A slightly slower migrating, non-specific band appears in all lanes, and this non-specific band was eliminated with monoclonal CA-specific Ab. Surprisingly, none of the GFP-tagged MMTV clones expresses CA protein. These results indicate that the GFP-tagged MMTVs are defective for both *sag* mRNA and CA production.

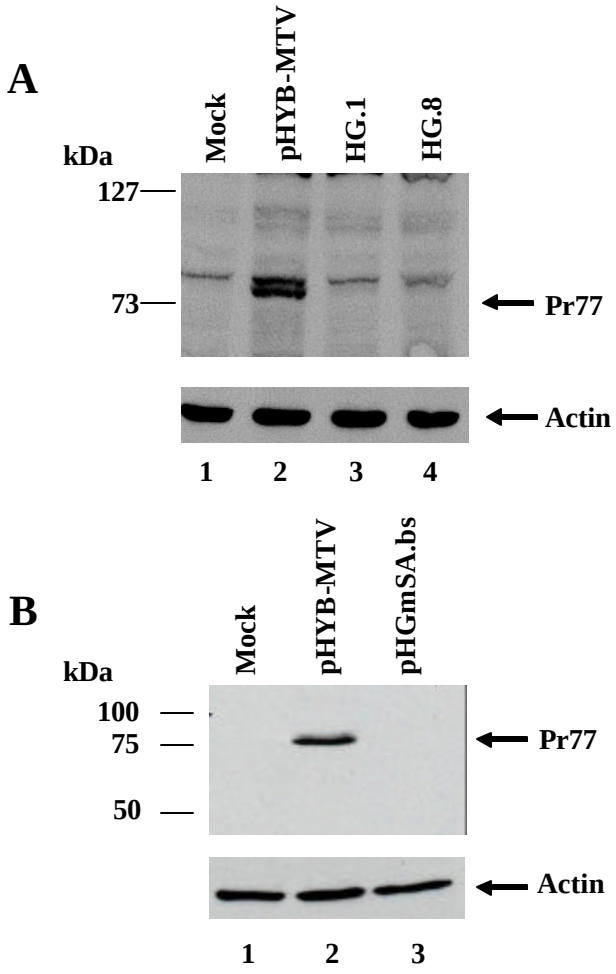


Fig. 3.21. Immunoblot analysis of CA expression from GFP-tagged MMTV mutants.

XC fibroblast cells were transiently transfected with either pHYB-MTV or the GFP-tagged MMTV mutants, and protein lysates were harvested at 48 hr post transfection. The proteins were separated using a 10% PAGE gel containing SDS, and (A) a polyclonal MMTV CA-specific antibody or (B) a monoclonal MMTV CA-specific antibody was used to detect the capsid precursor Pr77. The arrows indicate the expected products. Actin was used as a control for the amount of protein loaded.

3.2.6 Detection of unspliced mRNA in the cytoplasm of a GFP-tagged MMTV mutant

Spliced mRNAs are processed in the nucleus and exported to the cytoplasm via pathways mediated by different proteins, e.g., Crm1 or TAP (156,184,222). Synthesis of retroviral capsid proteins requires unspliced *gag* mRNA export from the nucleus for translation. Retroviruses use different strategies depending upon their complexity; however, the reported sequences needed for export are all located at the 3' end of the virus near the *env* gene (61,83,88,229). Since GFP-tagged MMTV clones all had mutations at the 3' end of the virus, it is possible that nuclear RNA export sequences also were altered.

The previous RT-PCR results were analyzed using total RNA. To determine whether the *gag* transcript could be exported to the cytoplasm, RNA was isolated from cytoplasmic and nuclear fractions of transiently transfected XC cells and subjected to RT-PCR analysis (Fig. 3.22). As expected, *gag* mRNA from pHYB-MTV was detected in both the nuclear and cytoplasmic fractions, although the cytoplasmic RNA level was much lower than the nuclear fraction. Surprisingly, the cytoplasmic fraction of pHYB-GFPmSA had undetectable *gag* mRNA, yet the *gag* transcript was observed in the nuclear fraction. The *env* transcript was detected in both cytoplasmic and nuclear fractions from pHYB-MTV and pHYB-GFPmSA transfected cells. These results suggest that CA

protein is not produced from the mutant provirus because the *gag* mRNA cannot be exported to the cytoplasm. Thus, the tagging of MMTV with GFP may have interfered with export of unspliced viral RNA.

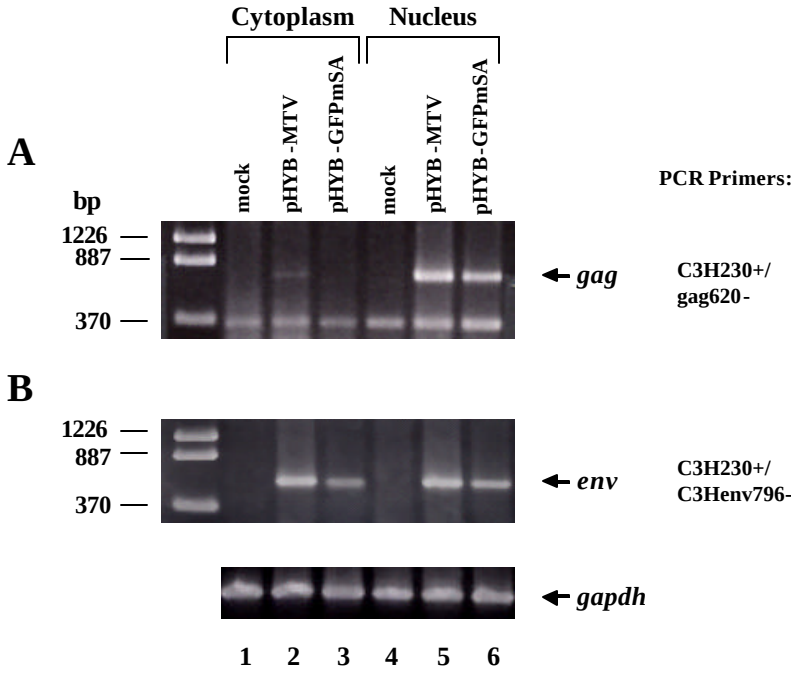


Fig. 3.22. Cytoplasmic and nuclear mRNA fractionation of XC cells transiently transfected with pHYB-MTV and GFP-tagged MMTV with a mutated splice acceptor.

XC fibroblast cells were transiently transfected with either pHYB-MTV or the GFP-tagged MMTV with a mutated splice acceptor, and cytoplasmic and nuclear RNA fractions were harvested at 48 hr post transfection. cDNA was prepared, and RT-PCR performed for (A) *gag* mRNA and (B) *env* mRNA using the indicated primers. Arrows show the locations of the expected products for each primer pair listed. The *gapdh* mRNA was used as a control, and all PCRs were analyzed by agarose gel electrophoresis.

3.3 DEVELOPMENT AND CHARACTERIZATION OF TRANSPOSON-MUTAGENIZED MMTVS

3.3.1 Production of transposon-mutagenized MMTVs

All retroviruses contain several important *cis*-acting elements (22,108,152,218), yet most elements have not been mapped in the MMTV provirus. The previous results suggested that the GFP-tagged provirus sustained alterations that prevented nuclear export of *gag-pol* mRNA. Development of a working vector for gene therapy requires identification of essential viral elements and, therefore, mutagenesis of pHYB-MTV was performed. The mutagenized pHYB-MTV was useful for localization of *cis*-acting elements but also for determining which genes and their products are necessary for MMTV trafficking to the mammary gland.

The pHYB-MTV plasmid was mutagenized *in vitro* using the DHFR-1 (trimethoprim) cassette as a transposon. The Epicentre EZ::TN transposon kit allowed for the transposon to be inserted randomly into either the MMTV provirus or the vector backbone, usually with one hit per plasmid. After electroporation into DH5α bacterial cells, individual colonies were screened by colony PCR with a panel of five primer pairs to identify the location of transposon insertion (Fig. 3.23). Screening continued until a pool of mutants for each region

was obtained. Several mutants from each region then were selected and analyzed by sequencing for the exact location of the transposon in pHYB-MTV (Fig. 3.24).

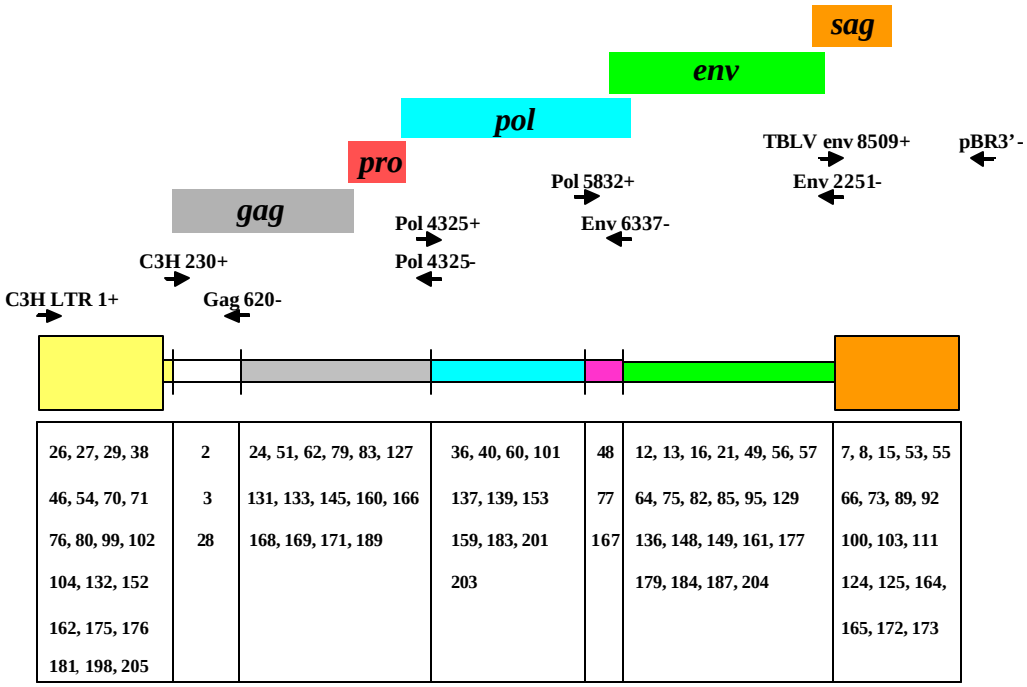


Fig. 3.23. Location of primers and trimethoprim cassettes within the mutagenized MMTV provirus.

The five primer pairs used for the colony PCR are shown above the MMTV provirus with arrows. The primer pairs divide the provirus into seven regions indicated by the different colors: yellow-5' LTR, white-*gag*, gray-*gag/pro*, blue-*pol*, pink- *pol/env*, green-*env*, and orange-3' LTR/*sag*. The pool of mutants obtained is shown in the box directly below each region.

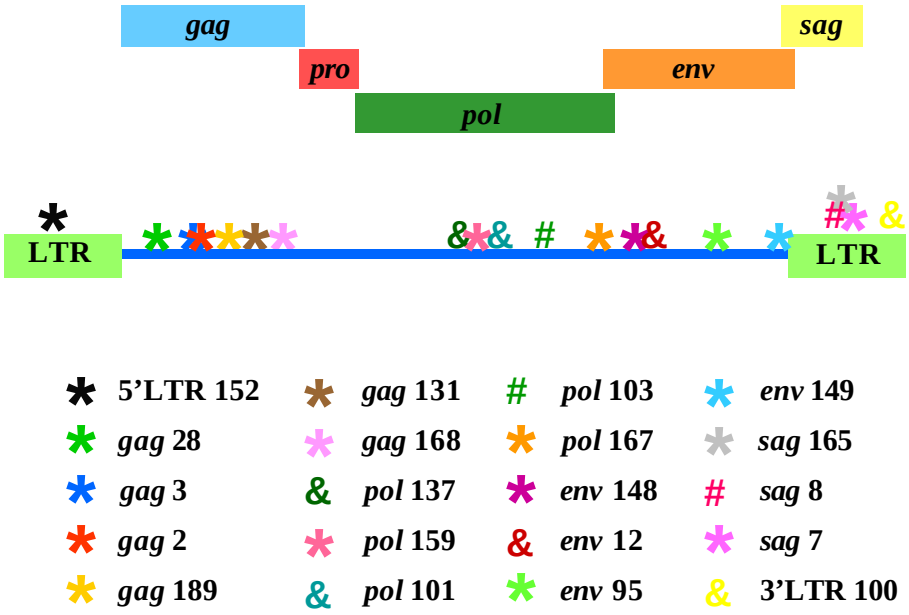


Fig. 3.24. Locations of transposon insertions in MMTV mutants.

The MMTV proteins are shown as colored boxes above the MMTV provirus. The asterisks, ampersands, and number signs indicate the location of transposon insertions that were localized by sequencing. Mutants were labeled according to the location of the insertion and the number of the clone.

3.3.2 Development and characterization of stably integrated transposon-mutagenized MMTVs

Previous experiments in the lab have shown that XC cells stably transfected with pHYB-MTV induce MMTV infection when injected into mice (155,253). To determine which viral genes are necessary for trafficking to the mammary gland, transposon mutants from different proviral regions were selected and stably transfected into XC cells. Retention of the stably integrated provirus and the transposon were verified by RT-PCR and immunoblot analysis.

Total RNA was harvested from the XC cells stably transfected with the mutated proviruses, and RT-PCR was performed to identify *gag*, *env*, and *sag* transcripts. Each mutant expressed bands of the same size as those produced by wild-type pHYB-MTV, unless the insertion occurred within that gene (Fig. 3.25 and 3.26). For example, the mutants *gag* 2 and *gag* 3 have the ca. 890 bp trimethoprim cassette inserted in the region between the primers used to detect *gag* mRNA and, therefore, due to the non-processivity of the polymerase used for the RT-PCR, a band is not produced. However, mutants *gag* 131 and *gag* 189 have an insertion outside the 230+ and gag620- primer pair, but within the *gag* gene and, thus, a band is observed for the *gag* mRNA (Fig 3.26A). Interestingly, some of the stably transfected XC cell lines did not show the RT-PCR expression pattern predicted from transposon location. For example, the mutant *gag* 28 does

not show a band for *gag* mRNA as predicted, but neither is *env* mRNA detected (Fig. 3.25B, lane 5) due to the transposon insertion just 5' to the *env* splice donor. The mutant *env* 95 also showed an unusual expression pattern for the *sag* mRNA (Fig. 3.26C, lane 6). A doublet was observed at the expected size of 723 bp, and another ~400 bp band was detected with this mutant, but not with other mutant-transfected cells or the positive control. The transposon insertion for mutant *env* 95 is not located near the SD or SA for *sag*, but it is possible that the insertion disrupted sequences necessary for proper splicing to occur.

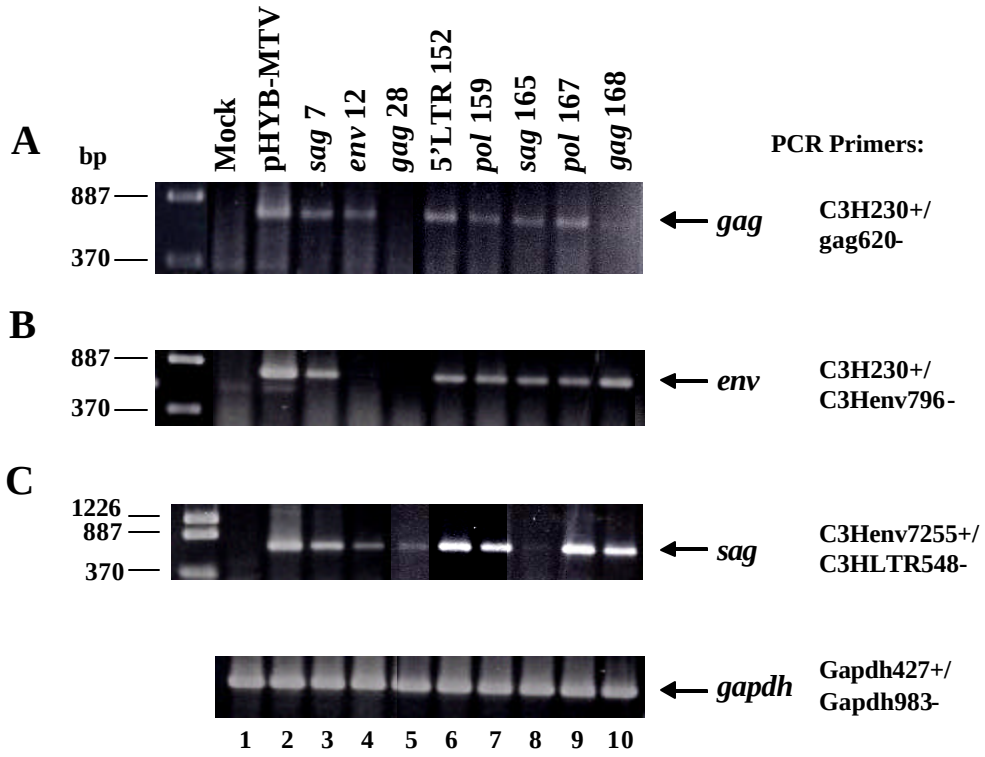


Fig. 3.25. Detection of mutant MMTV constructs in stably transfected XC cells.

XC fibroblast cells were stably transfected with transposon-mutagenized MMTVs. Total RNA was harvested, and cDNA was prepared for RT-PCR analysis of (A) *gag* mRNA, (B) *env* mRNA, and (C) *sag* mRNA using the indicated primers. The *gapdh* mRNA was used as a control, and PCRs were analyzed by agarose gel electrophoresis. Arrows show the location of the expected products for the primer pairs listed.

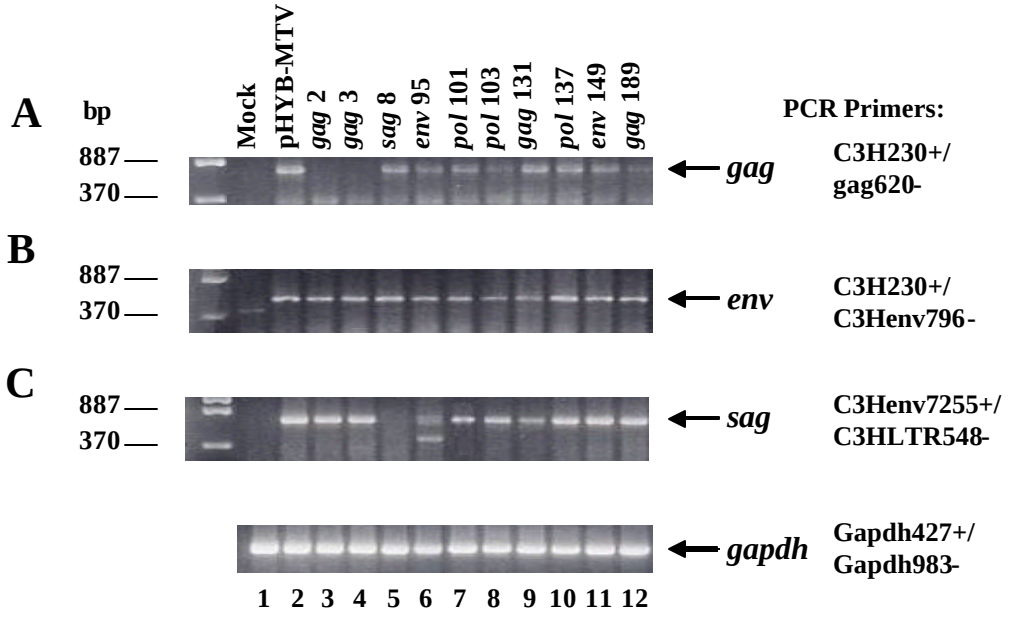


Fig. 3.26. Detection of mutant MMTV constructs in stably transfected XC cells.

XC fibroblast cells were stably transfected with transposon-mutagenized MMTVs. Total RNA was harvested, and cDNA was prepared for RT-PCR analysis of (A) *gag* mRNA, (B) *env* mRNA, and (C) *sag* mRNA using the indicated primers. Arrows show the location of the expected products for the primer pairs listed. The *gapdh* mRNA was used as a control, and all PCRs were analyzed by agarose gel electrophoresis.

Cell lysates also were harvested from the stably transfected mutant cell lines, and the monoclonal MMTV CA-specific Ab was used to detect CA protein using pHYB-MTV as a positive control (Fig. 3.27). As anticipated, the *pol*, *env*, *sag*, and LTR mutants all produced the capsid precursor Pr77, whereas none of the six *gag* mutants were able to produce any protein since the gene was disrupted by the transposon cassette. However, the protein levels of mutants *env* 12, *env* 95, and *sag* 8 were all significantly lower than wild type levels, despite the localization of each mutation outside the *gag* gene. The verification of mRNA and protein expression from these XC transfectants provides valuable reagents to assess *cis*-and *trans*-acting sequences required for MMTV replication and transmission.

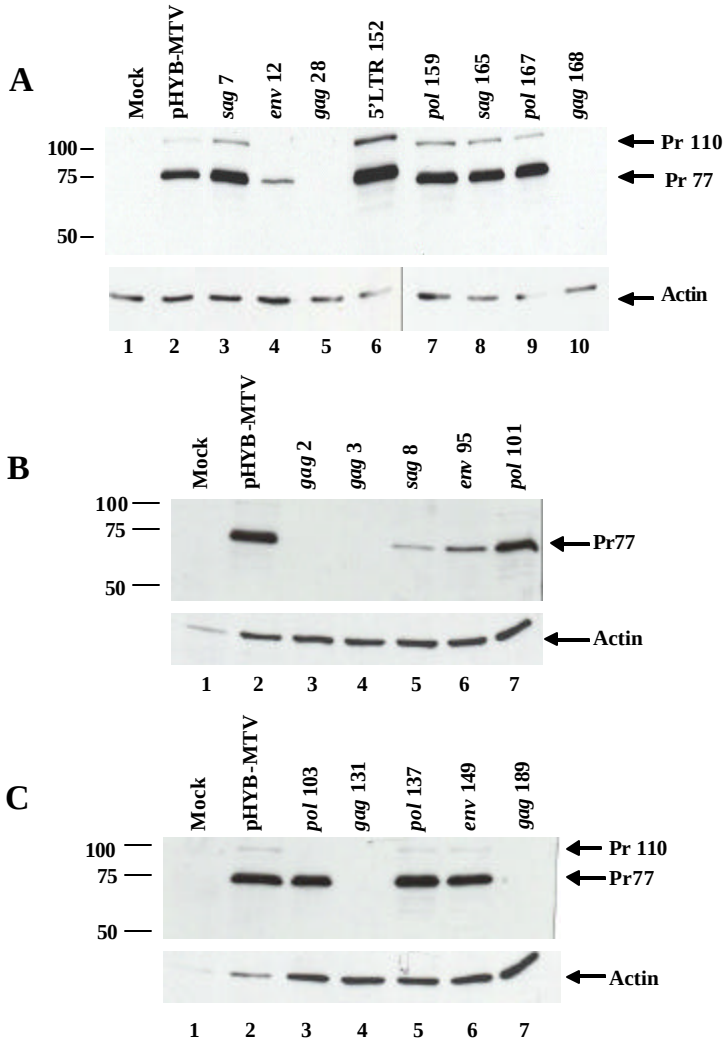


Fig. 3.27. Immunoblot analysis of CA expression from XC cells stably transfected with various MMTV constructs.

Protein lysates were harvested from XC cells stably transfected with wild-type or mutant MMTV proviruses and were separated on 10% PAGE gels containing SDS. Capsid precursors Pr110 and Pr77 were detected using an MMTV CA-specific antibody. The arrows indicate the expected products. Actin was used as a protein loading control.

3.3.3 Identification of MMTV genes necessary for trafficking to the mammary gland

After initial analysis of the stably transfected cell lines, mutants from each gene region were selected and injected into BALB/c weanling mice using the mutant 5' LTR 152 as a positive control. This mutant reverts to wild type viral RNA after DNA integration and transcription because an insertion in the 5' U3 region of the LTR is not transcribed by RNA polymerase II. The mice were analyzed two months after injection to monitor the percentage of CD4+Vβ14+ T cells, a measure of MMTV infection due to Sag-mediated deletion. The uninjected controls showed a normal percentage of CD4+Vβ14+ T cells, whereas the mutant 5' LTR 152 showed T-cell deletion that was statistically significant from the uninjected controls (Fig. 3.28). None of the transposon mutants deleted CD4+Vβ14+ T cells, which indicated that an active infection did not occur with any of the mutants. These results suggest that all of the viral proteins are needed for MMTV infection and subsequent T-cell deletion.

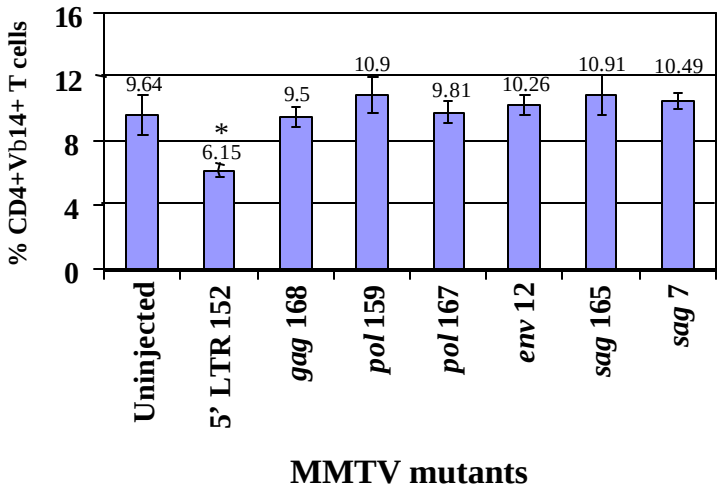


Fig. 3.28. CD4+Vβ14+ T cell deletion analysis from injected BALB/c mice.

XC cells stably transfected with the transposon-mutagenized MMTVs were injected into BALB/c mice and compared to uninjected mice. The mutant 5' LTR 152 was used as a positive control since wild type levels of viral proteins were detected despite the presence of a transposon in the 5' LTR (see text). The percentage of CD4+Vβ14+ T cells was determined using the FACS as described in the Materials and Methods. The values are shown as means of four injections with standard deviations. An asterisk indicates values that are significantly different from the uninjected control mice by the Student's t test (p=0.05).

The *sag* mutants theoretically lack Sag production, although this is difficult to verify because there are no available Sag-specific antibodies. Furthermore, these mutants would not delete the Vβ14+ T cells since Sag protein mediates this deletion. Therefore, MMTV transmission to the mammary gland was monitored by RT-PCR. Mammary gland RNA was harvested from each injected mouse (four mice for each construct) and analyzed for *gag* and *sag* mRNAs. The uninjected controls expressed neither C3H-specific *gag* or *sag* mRNA, as expected, and 75% of the positive controls injected with mutant 5' LTR 152 expressed both *gag* and *sag* mRNA (Fig. 3.29). However, RT-PCR analysis was unable to detect infection in mammary glands of one of the 5' LTR 152-injected mice, despite deletion of Vβ14+ T cells in this animal (Fig. 3.28). This result suggests that either the virus was not transmitted to the mammary gland or that the expression levels were too low to detect by RT-PCR. A *gag* mRNA product was not observed for any of the other mutant-injected mice, and most of these mice also showed no *sag* mRNA product. However, the *sag* 7 and *sag* 165 mutants each had one injected mouse that showed a *sag* mRNA product (Fig. 3.29 asterisks) even though a *gag* mRNA product was not detected for either mutant. It appears likely that the RT-PCR primers are able to detect *sag* mRNA levels more efficiently than the *gag* mRNA, which would explain detection of a *sag* mRNA product in the absence of *gag*. These results suggested that *sag*

mutants infect mammary glands with approximately 25% efficiency. These results also support previous findings from our lab, which showed that MMTV mutants expressing nonfunctional *sag* genes sporadically infect mammary glands of mice (155,253).

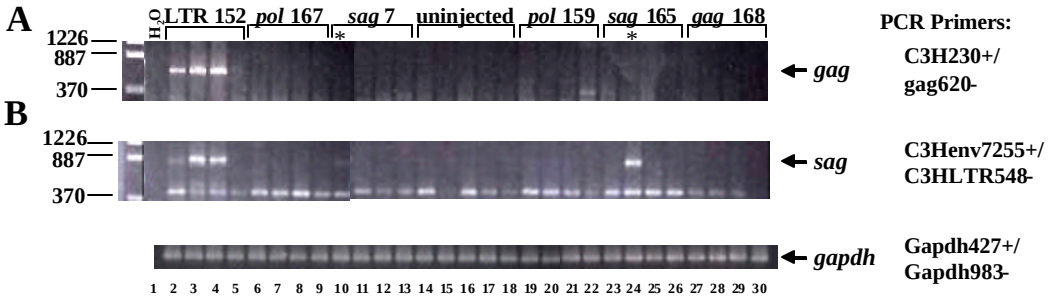


Fig. 3.29. Analysis of viral transcripts from mammary glands of injected BALB/c mice.

Stably transfected XC cells containing the transposon-mutagenized MMTVs were injected into BALB/c mice. Uninjected mice were used as negative controls and mice injected with the mutant 5' LTR 152 were used as positive controls. Total RNA was harvested from mammary gland tissue, and cDNA was prepared for RT-PCR analysis of (A) *gag* mRNA and (B) *sag* mRNA using the indicated primers. The arrows to the right of the gel show the products observed for each transcript. The *gapdh* mRNA was used as a control, and all PCRs were analyzed by agarose gel electrophoresis.

3.3.4 Identification of mutants unable to export unspliced mRNA to the cytoplasm

Previous results indicated that mutations in the 3' portion of pHYB-MTV compromise the expression of CA. To determine specific insertion sites that prevent CA expression, transient transfections of XC cells were performed using transposon mutants from the *pol*, *env*, *sag*, and the 3' LTR regions. Most of the mutants were able to express the CA precursor Pr77 at wild type levels; however, three of the *env* mutants did not express Pr77 at all (Fig. 3.30). Results from the transient transfections of some mutants did not correspond to the results from the stably transfected mutants. The mutants *sag* 8 and *env* 95 had reduced levels of Pr77 when stably integrated, whereas Pr77 expression appears to give wild-type levels after a transient transfection. The mutant *env* 149 had the exact opposite result since wild-type levels of Pr77 were observed in stable transfection assays, but no expression was detected in a transient transfection. Potentially, such different results could be due to: (i) loss of proviral expression from specific MMTV integration sites or cell clones in the infected population or (ii) the increased sensitivity of stable, compared to transient, assays due to the inefficiency of the latter.

RT-PCR analysis was performed on total RNA from the transient transfections to determine whether absence of CA expression was due to lack of

gag mRNA production. All transposon mutants expressed *gag* mRNA in transient assays (Fig. 3.31) and, therefore, should be able to make CA.

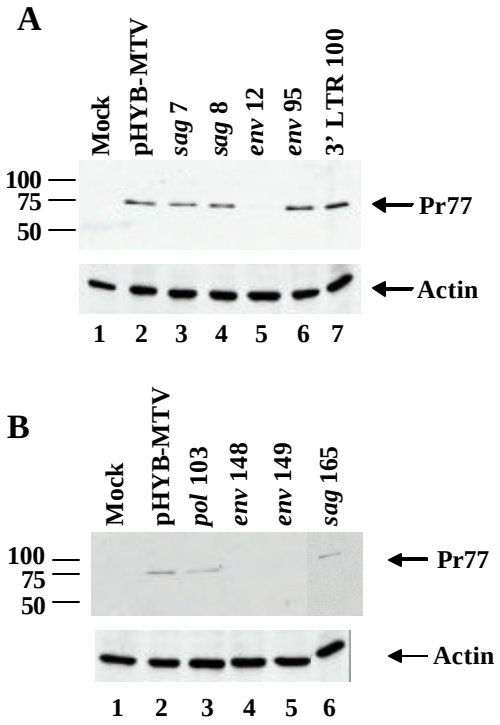


Fig. 3.30. Immunoblot analysis of CA expression from XC cells transiently transfected with transposon-mutagenized MMTVs.

XC fibroblast cells were transiently transfected with transposon-mutagenized MMTVs, and protein lysates were harvested. The proteins were separated using 10% PAGE gels containing SDS and detected using MMTV CA-specific antibody 48 hr post transfection. The arrows to the right of the gels indicate the expected products. Actin was used as a control for the amount of protein loaded.

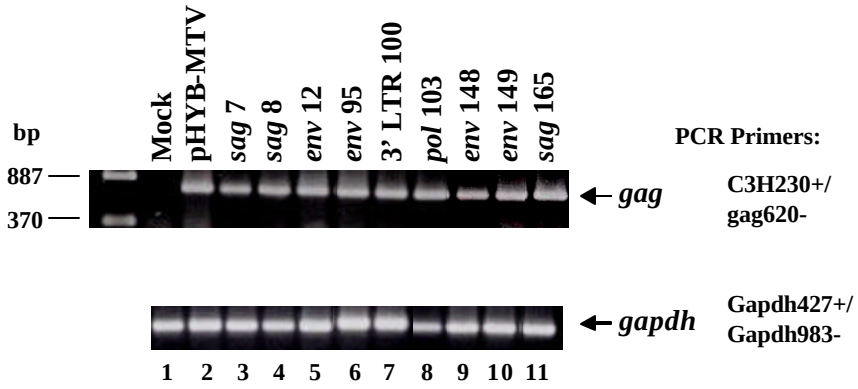


Fig. 3.31. Expression of *gag* mRNA in XC cells transiently transfected with transposon-mutagenized MMTVs.

XC fibroblast cells were transiently transfected with transposon-mutagenized MMTVs. Total RNA was harvested, and cDNA was prepared for RT-PCR analysis of *gag* mRNA using the indicated primers. Arrows show the location of the expected products for the primer pairs listed. The *gapdh* mRNA was used as a control, and all PCRs were analyzed by agarose gel electrophoresis.

Earlier experiments with the GFP-tagged MMTV showed that the *gag* mRNA, while expressed, was not exported to the cytoplasm. Therefore, the three *env* mutants that failed to produce Pr77 in transient transfections were subjected to RNA fractionation to determine the location of the *gag* mRNA. The results show that none of the mutants exported *gag* mRNA into the cytoplasm at wild-type levels, although low levels could be detected (Fig. 3.32). These results suggest that the *env* mutants are compromised for CA expression due to reduced export of the *gag* transcript. Two of the mutants, *env* 12 and *env* 148, are located approximately 200 bp apart at the 5' end of *env*, whereas mutant *env* 149 is located at the 3' end of *env* directly upstream of the SA for *sag* near the 3' LTR. This result suggested that there are two different elements or a bipartite element that is required for export of unspliced MMTV RNA to the cytoplasm.

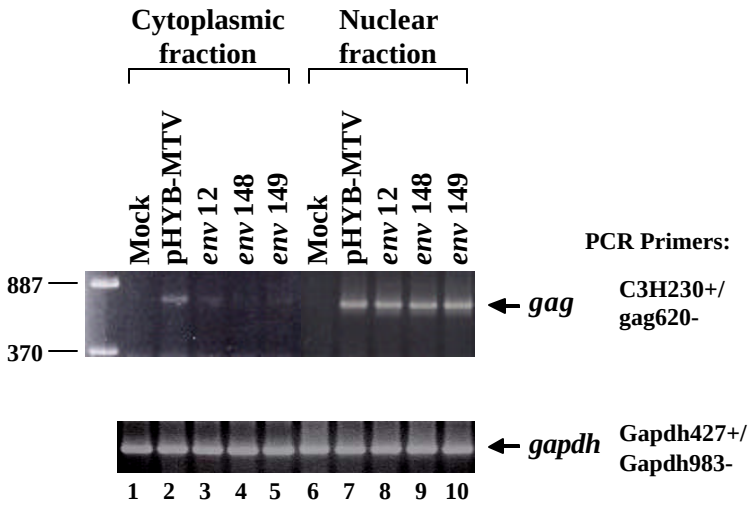


Fig. 3.32. Nuclear and cytoplasmic mRNA fractionation of XC cells transiently transfected with transposon-mutagenized MMTVs.

XC fibroblast cells were transiently transfected with three different *env* mutants, and cytoplasmic and nuclear RNA fractions were harvested 48 hr post transfection. cDNA was prepared, and RT-PCRs were performed using the indicated primers. The arrows show the location of the expected products for the primer pairs listed. The *gapdh* mRNA was used as a control, and all PCRs were analyzed by agarose gel electrophoresis.

3.4 IDENTIFICATION OF *CIS*- OR *TRANS*-ACTING ELEMENTS LOCATED IN THE MMTV PROVIRUS

3.4.1 Construction of a CTE-dependent reporter vector to characterize MMTV mRNA export

Retroviruses have different methods to export unspliced RNA to the cytoplasm depending on the complexity of the virus. Simple retroviruses, such as MPMV, depend on a CTE found in unspliced mRNAs as well as host factors that bind to the CTE for export of unspliced mRNAs to the cytoplasm (46,83,252). Complex retroviruses, such as HIV-1 and HTLV-1, encode proteins that bind to RNA elements located in unspliced or partially spliced mRNAs for cytoplasmic export (45,46,88). MMTV is classified as a simple retrovirus. Therefore, to determine whether MMTV contains a CTE, a CTE-dependent reporter gene was constructed.

The strategy used to construct the CTE-dependent reporter gene (Fig. 3.33) also has been described in the Materials and Methods. The *Renilla* luciferase reporter gene (*Rluc*) was utilized for the high sensitivity of the assay. Briefly, the chimeric intron from the vector pRL-TK was removed, and the SV40 late polyA signal was replaced with the bovine growth hormone (BGH) polyA signal to create pΔRL-BGH. pΔRL-BGH makes an unspliced *Rluc* mRNA, which is inefficiently exported into the cytoplasm unless a CTE is provided. To

determine CTE activity, MMTV fragments then were cloned into the *Not*I site using the MPMV CTE as a positive control. The MMTV fragments tested were derived from the 3' end of MMTV, and they encompass (i) the sites of all three *env* mutants, (ii) the two 5' *env* mutants, (iii) the one 3' *env* mutant, or (iv) the 3' MMTV LTR (Fig. 3.34). CTE activity is dependent upon orientation (251,252,256); thus, all fragments were cloned in both the forward (F) and reverse orientation (R).

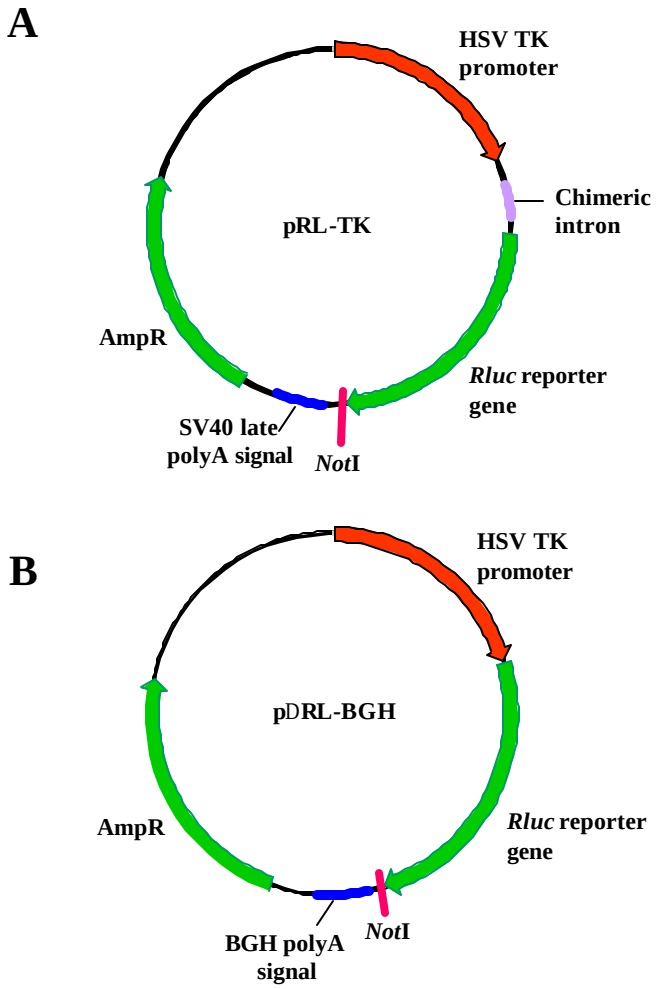


Fig. 3.33. Strategy to construct a CTE-dependent reporter gene.

A CTE-dependent luciferase vector was constructed using (A) the *Renilla* luciferase reporter vector pRL-TK. pRL-TK contains the *RLuc* reporter gene driven by the HSV TK promoter, a chimeric intron separating the gene from the promoter, an SV40 late polyA signal, and an ampicillin resistance gene. To create the CTE-dependent luciferase vector, pΔRL-BGH (B), the chimeric intron was removed and the SV40 polyA signal was replaced with the BGH polyA signal. MMTV fragments then were cloned into the *Not*I site for assays of CTE activity.

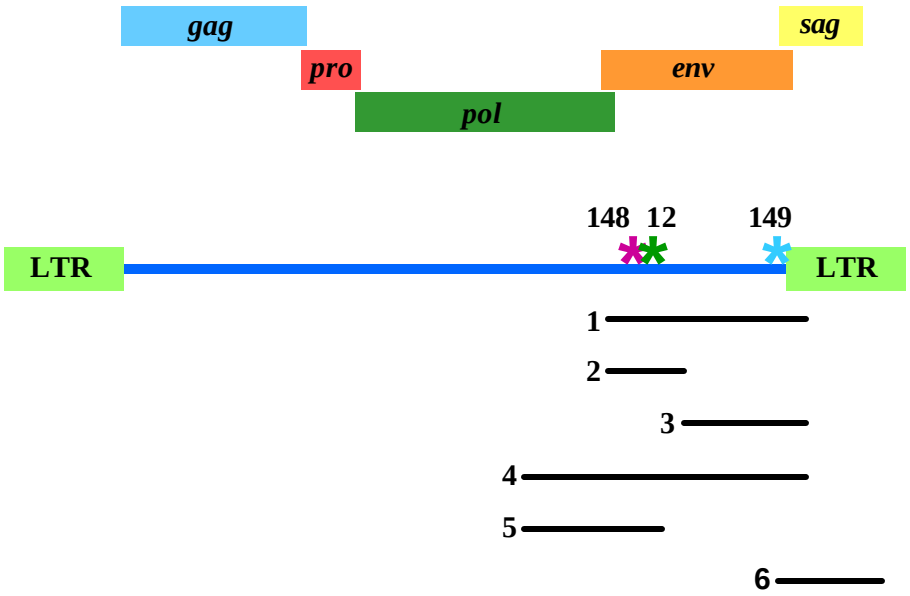


Fig. 3.34. Diagram of MMTV fragments used for assays of CTE activity.

The MMTV proteins are shown by colored boxes labeled with the protein above the MMTV provirus. The asterisks indicate the location of the transposon insertion for each *env* mutant unable to export unspliced mRNA to the cytoplasm. The plasmid pHYB-MTV was amplified by PCR with multiple primer sets to generate different sized MMTV fragments (black lines) for cloning into the pΔRL-BGH vector and analysis of CTE activity.

To determine which, if any, MMTV fragments contained a CTE, the newly constructed plasmids then were transiently transfected into XC cells. XC cells were used since they are able to express MMTV proteins, and CTE activity should be detectable in these cells. Cell lysates were harvested after 48 hr, and a luciferase assay was performed after normalization of *Renilla* luciferase (RLuc) activity for the percentage of cells positive for the co-transfected pEGFP-C1 (Fig. 3.35). The removal of the intron from the luciferase vector resulted in ca. 95-fold reduction in the relative RLuc activity. Addition of the MPMV CTE in the forward orientation showed a statistically significant increase over the background level of pΔRL-BGH; however, the increase was slightly less than 2-fold. The MPMV CTE in the reverse orientation had ~40% lower RLuc activity than the pΔRL-BGH vector. None of the MMTV fragments in the forward orientation showed an increase in RLuc activity over the pΔRL-BGH levels, but 2F and 6F had luciferase levels similar to pΔRL-BGH. MMTV fragments in the reverse orientation demonstrated 4- to 10-fold decreases in luciferase activity. These results suggested that the level of expression using the MPMV CTE (F), the positive control, was too low in XC cells for sensitive detection of CTE function. Therefore, it could not be determined if MMTV contains a CTE under these conditions.

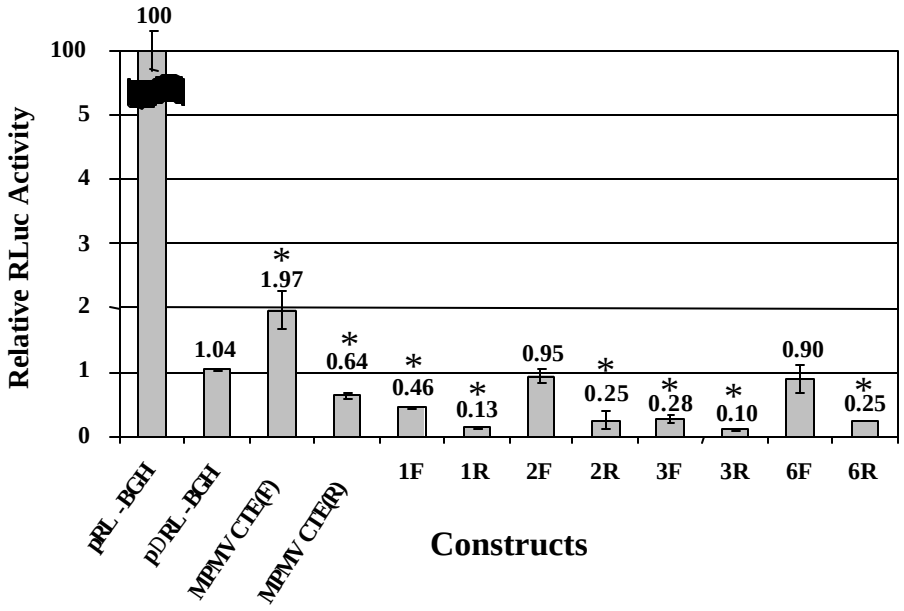


Fig. 3.35. Activity of various MMTV fragments for mRNA export in XC cells.

Luciferase reporter gene constructs were tested in XC cells. *Renilla* luciferase activity is given in light units/100 μg of protein normalized for DNA uptake as measured by co-transfection with the pEGFP-C1 plasmid. RLuc activity is reported relative to that of pRL-BGH, assigned a value of 100, and the relative RLuc activities are shown as means of triplicate assays with standard deviations. The positive control MPMV CTE forward orientation (F), the negative control MPMV CTE reverse orientation (R), and the MMTV fragments were compared to the activity of the pΔRL-BGH vector. An asterisk indicates values that are significantly different from the pΔRL-BGH vector control by the Student's t test (p=0.05).

Higher activity of the MPMV CTE has been reported in monkey and human cell lines (256). Therefore, 293T human kidney epithelial cells were selected for tests of CTE function because they also allow MMTV Gag production and, therefore, must have export of unspliced viral RNA. The 293T cells were transiently transfected, harvested, and assayed as previously described for the XC cells (Fig. 3.36). As expected, the removal of the intron from the luciferase vector resulted in ca. 110-fold decrease in RLuc activity. However, in 293T cells, the addition of the MPMV CTE in the forward orientation demonstrated ca. 3.5-fold increase in RLuc activity. This statistically significant increase is similar to the level reported with other CTE-dependent reporter vectors (164,256). The MMTV fragments in the forward orientation again showed no increase in RLuc activity over levels observed with pΔRL-BGH. All fragments, including the MPMV CTE, in the reverse orientation showed a decrease in activity from pΔRL-BGH levels as previously seen in the XC cells. This experiment also was repeated in Jurkat human T cells and A20 mouse B cells, which are able to express MMTV Gag proteins, with similar results (data not shown). The MPMV CTE system, while inefficient in human cells, clearly showed a significant difference over background compared to results from rat (XC) or mouse (A20) cells, and none of the MMTV fragments in the forward

orientation had increased RLuc activity over background pΔRL-BGH levels. Together, these results suggest that MMTV lacks a CTE.

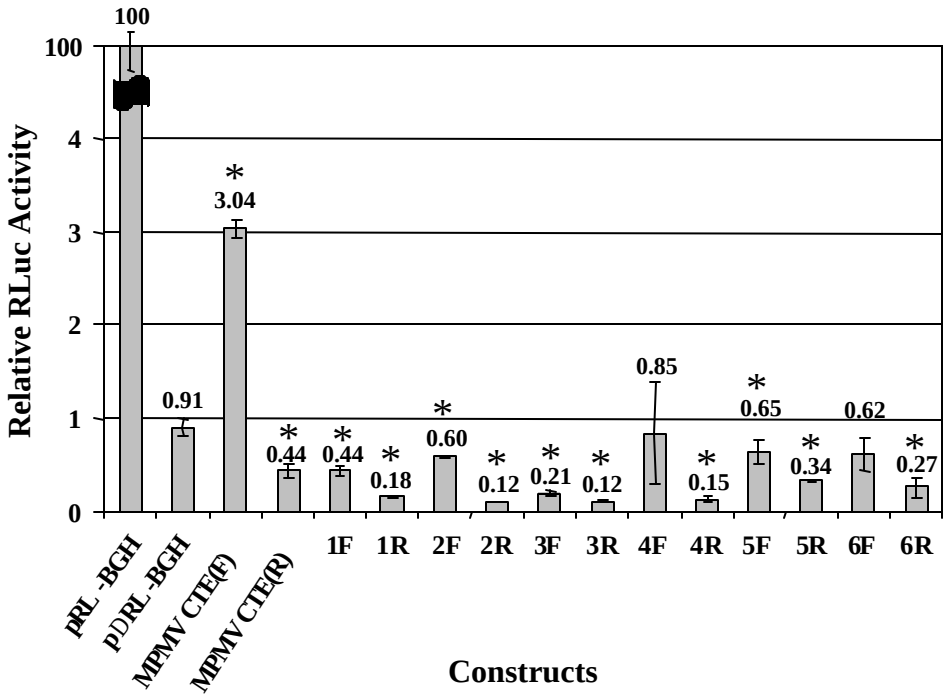


Fig. 3.36. Activity of various MMTV fragments for mRNA export in 293T cells.

Renilla luciferase activity is reported as described in Fig. 3.35. An asterisk indicates values that are significantly different from the pΔRL-BGH vector control by the Student's t test (p=0.05).

3.4.2 Detection of novel MMTV mRNAs

MMTV requires export of unspliced and partially spliced mRNAs to the cytoplasm, so in the absence of a CTE, the virus must use an alternative mRNA export pathway. The human endogenous retroviruses type K (HERV-K) are members of the betaretroviruses, which include MMTV, and HERV-K proviruses recently have been shown to produce a doubly spliced mRNA encoding an RNA export protein, Rec (130,131,134). MMTV produces Gag, Pol, Env, and Sag proteins from their corresponding mRNAs (92,105,155), and except for the unspliced *gag-pol* mRNA, all other known mRNAs are singly spliced (Fig. 3.37). Therefore to detect novel MMTV mRNAs, RT-PCR was performed on total RNA harvested from XC cells stably transfected with pHYB-MTV. Because all MMTV strains should require mRNA export, Jurkat cells stably transfected with the MMTV strain TBLV also were used for RT-PCR. MMTV and TBLV are 98% identical with the differences occurring only in the LTRs. Since the rest of the proviruses are identical, novel mRNAs detected in MMTV should also be detected in TBLV. The primer pairs C3H 230+/C3H LTR 548- and C3H 230+/C3H LTR 408- are at opposite ends of the provirus and were used to detect novel spliced mRNAs. The primer C3H LTR 408- was used to detect mRNAs in TBLV because the primer C3H LTR 548- is located in the negative regulatory elements, a region deleted from the TBLV LTRs (10,32,50). As expected, both

stable cell lines expressed both *env* and *sag* mRNAs (Fig. 3.38). However, another strong band was detected between these mRNAs (thick arrow). The novel band from XC and Jurkat stably transfected cells was isolated, and sequence analysis verified that this prominent new band was derived from a doubly spliced transcript. Several other minor bands also were observed.

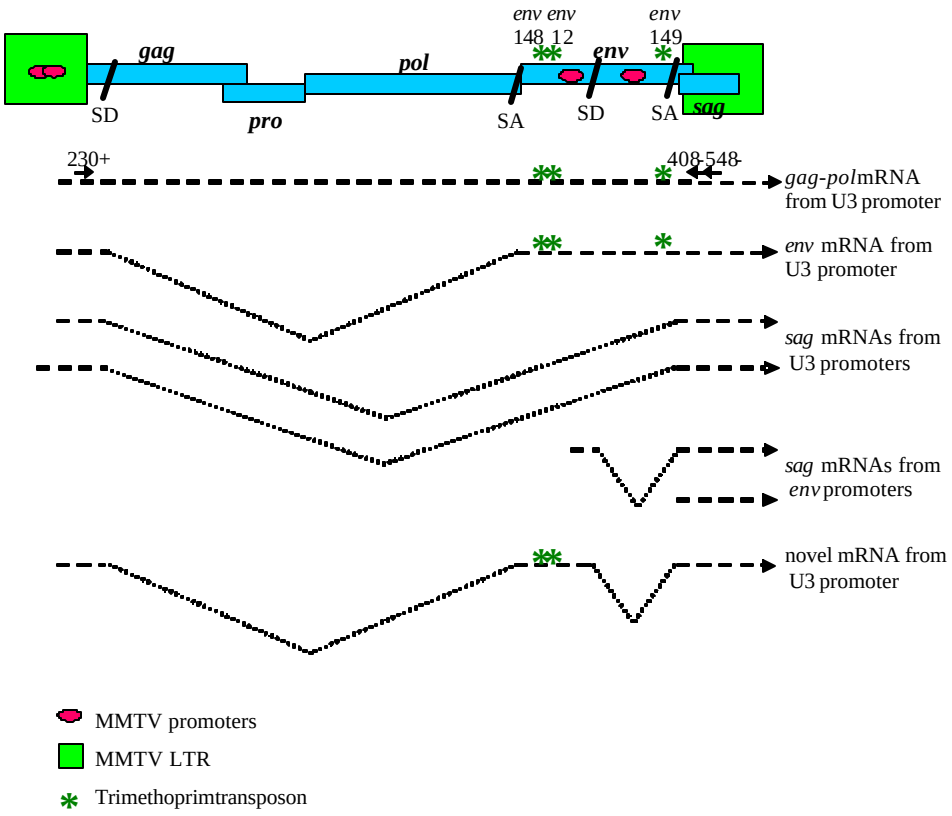


Fig. 3.37. Diagram of the known and newly identified MMTV mRNAs.

The MMTV provirus is shown with the LTRs as green boxes and the proteins as blue rectangles. MMTV promoters are represented by pink ovals. The mRNA transcripts expressed by MMTV are shown below the provirus as dashed lines. Splice donors (SD) and splice acceptors (SA) are indicated by black lines, and introns are shown as dotted lines with a "V" shape. The location of the trimethoprim insertion for each labeled mutant is represented by a dark green asterisk. Primers used to detect novel mRNAs are indicated above the *gag-pol* mRNA [Adapted from Mustafa *et al*, 2000 (155)].

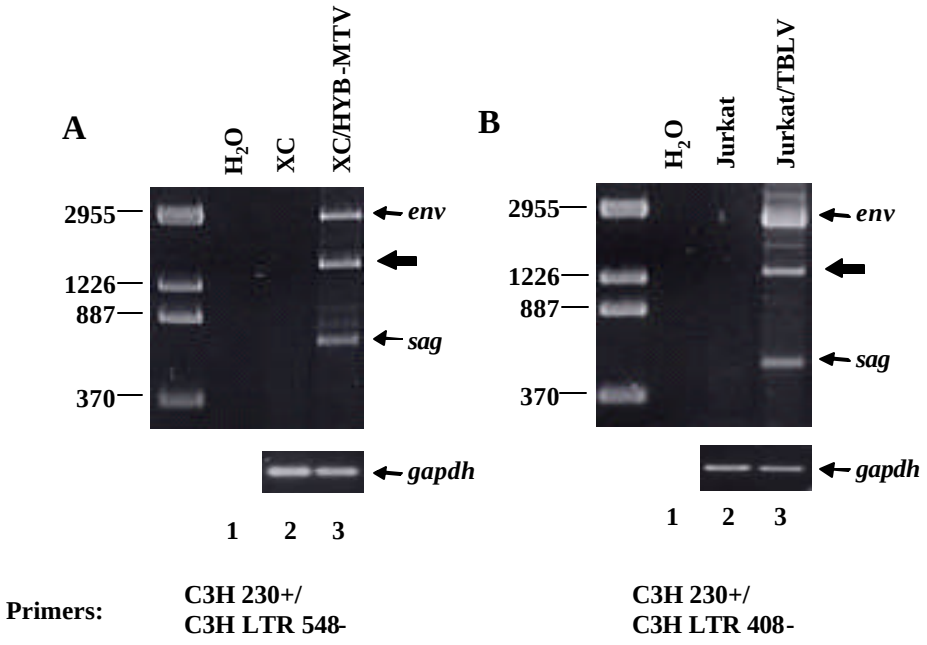


Fig. 3.38. Detection of novel viral transcripts in XC and Jurkat stable cell lines.

Total RNA was harvested from (A) XC fibroblast cells and (B) Jurkat T cells stably transfected with pHYB-MTV and pTBLV, respectively, at 48 hr post Dex induction. cDNA was prepared, and RT-PCR was performed using the indicated primers. Arrows show the location of the transcripts for each primer pair listed. The *gapdh* mRNA was used as a control, and all PCRs were analyzed by agarose gel electrophoresis.

The sequence of the novel 1500 bp RT-PCR product then was translated into an open reading frame of 301 amino acids. This putative protein contained the Env leader peptide, or signal sequence, 162 amino acids from gp52 (SU), and 41 amino acids from gp36 (TM) (Fig. 3.39). Predicted protein motifs found using Motif Scan and ExPASy proteomics tools showed common motifs such as glycosylation, myristoylation, amidation, and protein kinase C phosphorylation sites. Interestingly, this putative product also contained an NLS and an NES, which closely resemble sequences from Rev-related proteins (Figs. 3.39 and 3.40), as well as an arginine-rich motif (ARM), which is an RNA-binding domain found in the Rev-related proteins (174). These results suggest that MMTV contains a *trans*-acting protein used to export unspliced and partially spliced mRNAs to the cytoplasm. This putative protein was named Rem for regulator of export of MMTV.

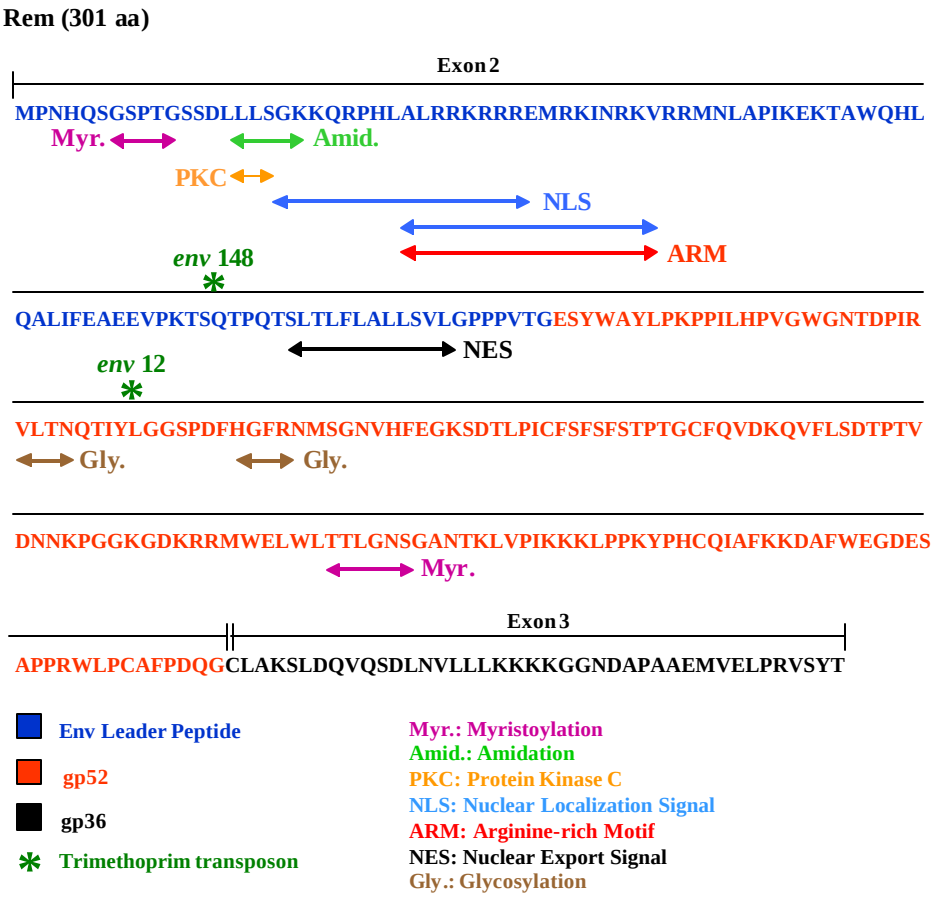


Fig. 3.39. Predicted motifs of the putative MMTV-encoded protein Rem.

The putative protein Rem is shown as an amino acid sequence. The Env leader peptide is indicated by blue letters, the SU sequence is indicated by red letters, and the TM sequence is indicated by black letters. The exons are shown as lines above the amino acid sequence, and the predicted motifs are labeled below the sequence and indicated by horizontal arrows. The location of the trimethoprim insertion for each labeled mutant is represented by a dark green asterisk.

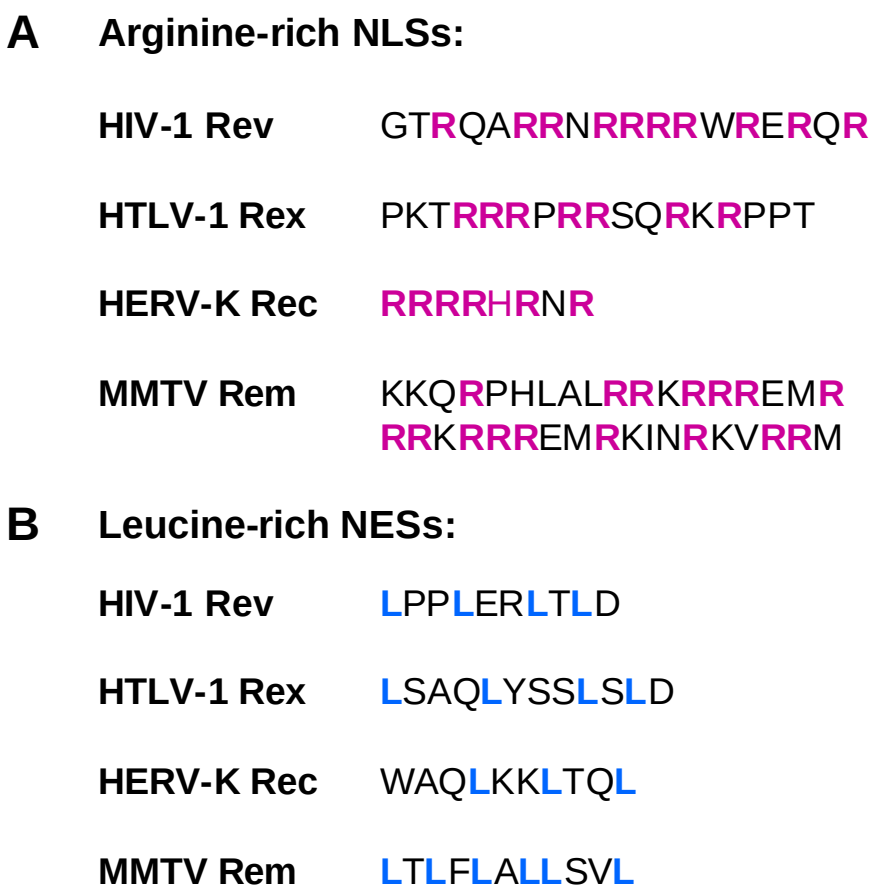


Fig. 3.40. Comparison of the NLS and NES sequences of Rev-related proteins.

(A) The potential NLS of MMTV Rem is compared to Rev and other Rev-like proteins. The arginine-rich sequence is represented by purple arginines (**R**). (B) The potential NES of MMTV Rem is compared to Rev and other Rev-like proteins. The leucine-rich sequence is indicated by blue leucines (**L**).

3.4.3 Complementation analysis to detect an MMTV-encoded *trans*-acting factor

Complex retroviruses produce *trans*-acting proteins, such as Rev, which bind to viral RNAs and a cellular export factor, such as Crm1 (88,174). Two of the three *env* mutants that were unable to transport unspliced mRNA to the cytoplasm had trimethoprim cassette insertions, one in the Env leader sequence and one in SU, in the Rem open reading frame. To determine if MMTV expresses a *trans*-acting factor, complementation experiments were performed (see Fig. 3.41 for strategy). Briefly, cells stably transfected with the *env* mutant only allowed transport of completely spliced mRNAs to the cytoplasm and, therefore, unspliced or partially spliced mRNAs can not be detected in the cytoplasm. If MMTV expresses a *trans*-acting RNA export factor, then transient transfections of pHYB-MTV into the *env* mutant stable cells should allow export of the unspliced or partially spliced mRNAs into the cytoplasm if the *env* mutant transcripts contain an RRE-like element. XC/*env* mutant stable cells were transiently transfected with pHYB-MTV. Similarly, XC cells stably transfected with pHYB-MTV were transiently transfected with the *env* mutant plasmids, which also should allow export of unspliced or partially spliced mRNAs. RNA was isolated from cytoplasmic and nuclear fractions of these cells as well as untransfected cells and subjected to RT-PCR analysis (Fig. 3.42). Primers in *pol*

and the trimethoprim cassette were used to detect only the mutant unspliced mRNA. As expected, the unspliced mRNA was detectable in the nuclear fraction of both untransfected and transfected mutant *env* 12 and 148 stable cells, but not in the untransfected cytoplasmic fraction of the mutant *env* 12 and 148 stable cells. However, when pHYB-MTV was transfected into XC cells stably transfected with mutant *env* 148, or the mutant *env* 148 plasmid was transfected into XC cells stably transfected with pHYB-MTV, unspliced mRNA was detectable in cytoplasmic fractions (purple asterisks).

The recovery of the unspliced mRNA was only observed with the mutant *env* 148 and not the mutant *env* 12. One possible explanation is that the insertion in the mutant *env* 12 disrupted both a *cis*-acting element in the RNA and the Rem open reading frame. Disruption of the RNA element would prevent complementation even in the presence of the *trans*-acting protein. The *sag* mRNA was used as a control for export to the cytoplasm, because this RNA is completely spliced and able to be exported by normal mechanisms. As expected, the *sag* transcript was detected in the cytoplasmic fraction as well as the nuclear fraction for all transfected samples. These results indicate that MMTV expresses a *trans*-acting factor that functions in export of unspliced or partially spliced mRNAs and that the *env* 148 mutation inactivates the export function of this factor.

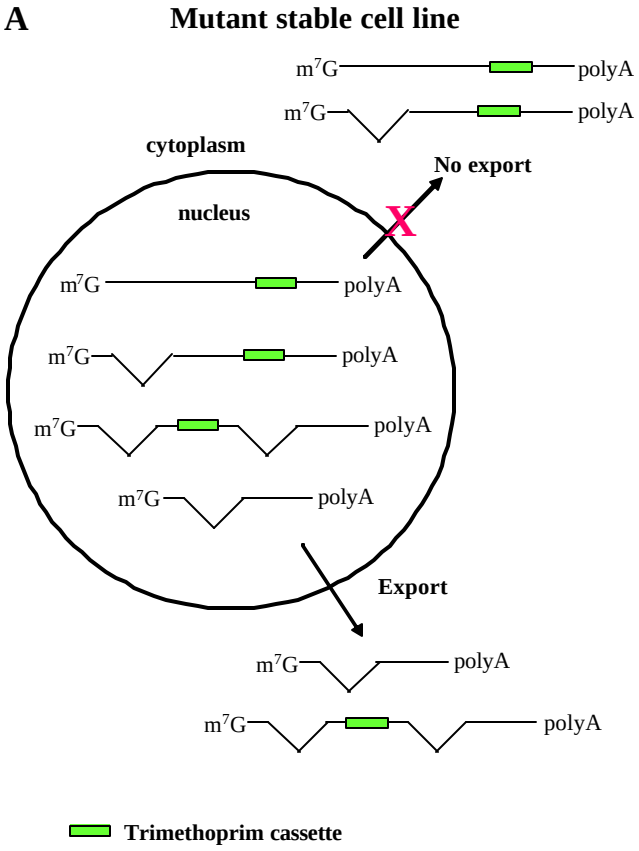
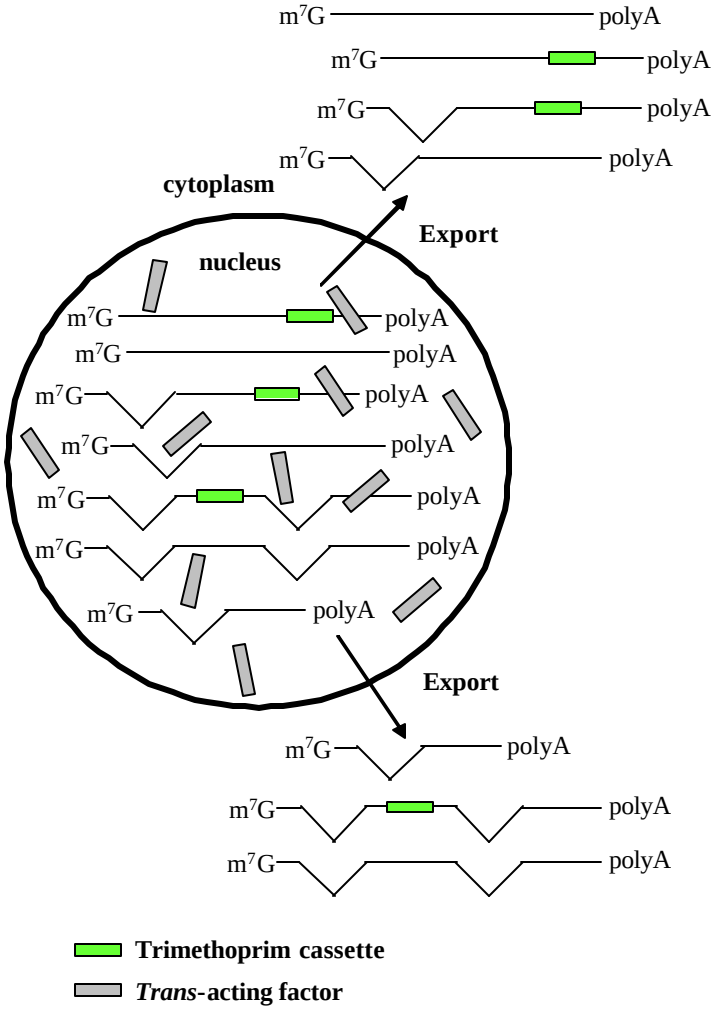


Fig. 3.41. Strategy for the complementation of the *env* mutant stable cell lines with the wild-type MMTV plasmid.

(A) The cytoplasmic and nuclear fractions of mRNA transcripts from the *env* mutant stable cell line are shown. All transcripts are indicated by a straight line flanked by a cap (m^7G) and a polyA with some containing introns shown as lines in a "V" shape and the trimethoprim cassette indicated as a green rectangle. All transcripts are expressed in the nucleus, but only the completely spliced mRNAs are exported. The unspliced or partially spliced mRNAs are not exported to the cytoplasm indicated by the red X. (B) The possible cytoplasmic and nuclear fractions of mRNA transcripts from the *env* mutant stable cell line and the transiently transfected pHYB-MTV are shown. The potential MMTV *trans*-acting factor is indicated by a gray rectangle and causes export of the unspliced and partially spliced mRNAs of the *env* mutant.

B

Mutant stable cell line transfected with pHYB-MTV



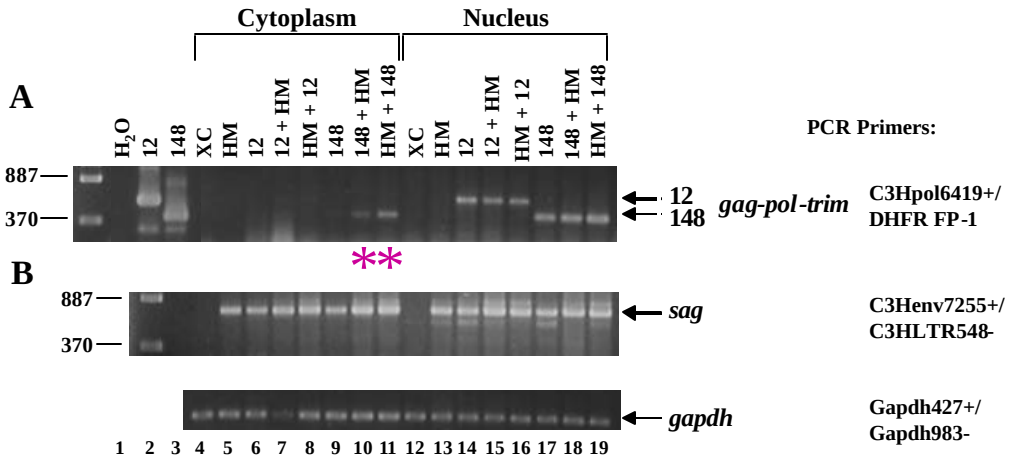


Fig. 3.42. Nuclear and cytoplasmic mRNA fractionation of transiently transfected XC stable cell lines.

XC cells stably transfected with either pHYB-MTV (HM) or *env* mutants (12 or 148) were transiently transfected respectively with the *env* mutant plasmids or pHYB-MTV, and cytoplasmic and nuclear RNA fractions were harvested at 48 hr post transfection. cDNA was prepared, and RT-PCR performed for (A) *gag-pol-trim* mRNA or (B) *sag* mRNA using the indicated primers. Arrows show the location of the expected products for each primer pair listed, and the purple asterisks indicate complementation occurring. The *gapdh* mRNA was used as a control, and all PCRs were analyzed by agarose gel electrophoresis.

4. Discussion

4.1 DEVELOPMENT OF A BACTERIAL SYSTEM FOR GENE DELIVERY TO MAMMALIAN CELLS

Gene therapy requires both targeting of specific cells and efficient delivery and expression of genes to these cells (96,197,240,249,259). Most gene therapy vectors accomplish at least one of the requirements; however, an ideal vector would be able to achieve all requirements. Retroviral vectors are able to integrate into the host cell genome allowing stable and sustained expression of the delivered gene, yet many retroviruses infect a broad range of host cells instead of targeting specific cells (114,116,165). MMTV infects lymphoid cells in the GALT after translocation through M cells, which then deliver the virus to mammary epithelial cells for subsequent infection and replication (56,72,142,180). The targeting of mammary gland epithelial cells for infection and ability to integrate into the chromosome make MMTV an ideal vector for breast cancer gene therapy, yet MMTV infection of human cells appears to be inefficient. Therefore a specific and efficient system for the delivery of MMTV vectors to the M cells and GALT is needed.

Bacterial delivery systems have been shown to deliver DNA to eukaryotic cells, and many of the bacterial systems will target M cells and the GALT *in vivo* (245,246). The combination of viral and bacterial vectors would create one gene

therapy system that consists of (i) bacteria containing the MMTV-based vector that traffic to M cells and deliver the provirus to lymphoid cells and (ii) an MMTV provirus capable of trafficking from lymphocytes to mammary epithelial cells where expression of the delivered gene can occur.

4.1.1 Delivery by *Shigella* to lymphoid cells is inefficient

Bacterial delivery of DNA to eukaryotic cells is usually accomplished through attenuated bacteria that escape from the phagocytic vacuole and lyse to release the DNA into the cytoplasm of the host. *Shigella* require escape from a vacuole for replication, and several attenuated vectors that lyse in eukaryotic cells have been developed (43,160,161,245,246). Previous studies have shown that a *Shigella asd* mutant, which requires the prokaryotic amino acid DAP, invades eukaryotic cells but will lyse within the cytoplasm due to the absence of DAP. This *asd* mutant also was shown to deliver both β-galactosidase and GFP to baby hamster kidney cells and human epithelial cells, respectively (207)(S. Seliger, personal communication). However, DNA delivery to lymphoid cells is necessary for trafficking of the MMTV provirus to the mammary gland.

Human and mouse lymphoid cells were infected with several *Shigella asd* mutants containing GFP vectors driven by various promoters. Delivery of the GFP vectors to lymphoid cells was not observed since significant GFP expression

with any of the vectors was not detected. Comparisons of GFP expression in human Jurkat T cells between a transient transfection and bacterial infection indicated that the transient transfection was 50-fold more efficient than the bacterial delivery method (Fig. 3.1). PCR analysis of the cytoplasmic and nuclear DNA fractions from both the transient transfection and bacterial infection also confirmed that delivery of GFP vectors from the *Shigella asd* mutant was inefficient when compared to transient transfection methods such as Superfect.

The differences between the previous results and the results shown here may be due to use of either adherent or suspension cells in the assay. Adherent epithelial cells, which were used successfully for gene delivery, were grown in a monolayer that allowed each cell to contact the bacteria added. However, lymphoid cells were grown in suspension and required a modified infection protocol involving the centrifugation of the bacteria onto the cells in a conical tube. This modification may have resulted in two layers, with the bacteria overlaying the lymphoid cell layer. If this is the case, only those cells at the top of the layer will contact the bacteria and decrease the efficiency of delivery. Attempts to alter the bacterial infection and increase the efficiency of delivery were unsuccessful.

In vivo delivery of GFP by the *Shigella asd* mutant to B and T cells in the GALT also was performed, but expression of GFP in these lymphoid cells was

not observed. Instead, GFP expression was detected in the M cells and, therefore, lysis of the *Shigella asd* mutant apparently occurred prematurely in these cells (S. Seliger, personal communication). These results showed the inefficiency of the *Shigella asd* mutant as a delivery system to lymphoid cells in the GALT and prompted a change to an alternative system.

4.1.2 Gene delivery by *Yersinia* causes toxicity to eukaryotic cells

Most bacterial delivery systems require lysis of the bacteria for delivery of a gene to the cytoplasm of a host cell; however, previous results using a *Shigella asd* mutant have indicated that lysis occurs too early for DNA delivery to lymphoid cells in the GALT. Therefore, a bacterial system that did not require lysis for delivery was developed using the *Yersinia* TTSS. *Yersinia* replicate extracellularly in the GALT, and the TTSS allows the bacteria to evade the immune system by binding to immune cells and translocating proteins, or Yops, into the cytoplasm of the host cell.

The TTSS usually delivers proteins to the cytoplasm of eukaryotic cells, but the delivery system for gene therapy requires DNA delivery to the nucleus. Only the Yop secretion and translocation signals are needed to translocate other proteins into mammalian cells and, therefore, it was predicted that a single protein could bind DNA and allow translocation and delivery of both the protein and

DNA into the cytoplasm of a host cell (Fig. 3.3). A fusion protein containing the secretion and translocation signals of the YopE protein and the *lac* repressor (*lacI*) was constructed in addition to a GFP vector containing *lac* operators to which *lacI* could bind. Both were transformed into a *Yersinia* multiple Yop mutant (MYM) for use in delivery assays. Delivery of GFP to human epithelial cells using this method has been shown by Dr. Shelley Payne's lab (personal communication).

Mouse and human lymphoid cells were infected with the *Yersinia* MYM delivery system, yet no detectable GFP expression was observed. Surprisingly, the addition of *Yersinia* to these eukaryotic cells caused an unusually high amount of cell death compared to other bacterial infections analyzed after 48 hr. Cell toxicity was not reported for previous experiments using the fusion protein containing adenylate cyclase; however, *Yersinia* were removed from the cultured cells after only two hr, and the cells then were harvested for cAMP levels (19,98,211). The short incubation of the *Yersinia* with the eukaryotic cells may not have detected toxic effects. Previous results of GFP delivery to human epithelial cells also did not demonstrate cell toxicity (personal communication), although our results using this assay resulted in detachment of the epithelial cells, indicative of toxicity. FACS analysis of the epithelial cells at different times after infection with *Yersinia* indicated that toxicity started after 24 hr and continued through 48 hr when compared to mock-infected cells. GFP expression from these

Yersinia-infected epithelial cells also was not observed after gating on the normal population seen in the mock-infected cells. However, a peak appeared when the entire population, including the dispersed and dying cells, was analyzed, suggesting autofluorescence from dead cells (Fig. 3.7B).

Together, it appears that *Yersinia* infection of cultured eukaryotic cells, both adherent and suspension cultures, causes cell toxicity. Such results are not unexpected from wild-type *Yersinia* since the Yop proteins translocated into the cytoplasm of host cells exert their effects on cellular pathways, including those leading to apoptosis. YopE, YopH and YopO have all been shown to have cytotoxic effects on cultured epithelial cells (42,82,192,193) and, therefore, the *Yersinia* strain utilized in these infections was the MYM strain, which contains mutations in *yopH, E, K, O,* and *M*. After the development of the *Yersinia* MYM strain, a new protein YopT was identified that also has a cytotoxic effect on epithelial cells (42,98). This protein is expressed from the *Yersinia* MYM strain and explains the cytotoxic effects observed in our experiments. The effect of YopT on lymphoid cells has not been reported, although the cytotoxicity detected in the MYM infections of lymphoid cells indicates the same effect as on the epithelial cells. Deletion of YopT from the MYM strain was not performed since YopT was identified recently, and it is possible that other unidentified *Yersinia* proteins also cause eukaryotic cell death. Thus, the relatively uncharacterized

genetic background of *Yersinia* prompted use of a more defined genetic system for bacterial gene delivery.

4.1.3 *E. coli* carrying *invA* are unable to specifically target lymphoid cells

The *Yersinia* delivery system appeared to cause cell toxicity. Wild type *E. coli* is not invasive, but the genome is well characterized and can easily be manipulated. Therefore, a delivery system was developed using *E. coli* containing the *Yersinia* TTSS. The invasin gene (*invA*) from *Yersinia enterocolitica* also was expressed in *E. coli* to allow attachment to mammalian cells expressing β1 integrins (16,21,100,101). Since lymphoid cells express β1 integrins (8,171,248,258), *E. coli* carrying the *invA* and TTSS plasmids should target the lymphoid cells in the Peyer's patches and deliver DNA into their cytoplasm. The plasmid encoding the TTSS, pIB29MEK, contains mutations in *yopH* and *yopE*, but not *yopO* or *yopT* which are all known to cause cell toxicity. Thus infections of eukaryotic cells with *E. coli* containing pIB29MEK and another plasmid known to not cause cytotoxicity, pHYB-MTV, were performed.

FACS analysis of an epithelial cell line and a B cell line indicated that *E. coli* containing the *Yersinia* TTSS did not cause cell toxicity after 48 hr. It is possible that YopO and YopT were not expressed in sufficient quantities to have an effect on cells or, alternatively, a protein expressed from the *Yersinia*

chromosome is needed for the regulation of these proteins. Although eukaryotic cell lysis could be prevented using a different bacterial system, it appears that *E. coli* were unable to attach to the eukaryotic cells using the invasin provided from *Yersinia*. Lack of cell attachment would prevent translocation of the Yops into the cytoplasm of the host cell and, therefore, no cell toxicity would be observed.

To assess attachment, *E. coli* containing both the invasin gene and a prokaryotic GFP plasmid together with control *E. coli* that did not contain the invasin gene were utilized in infections of eukaryotic cells known to express β1 integrins. Only *E. coli* with the invasin gene should bind these cells and display fluorescence due to the GFP. *E. coli* without invasin should be unable to bind specifically, and the eukaryotic cells then would have the same fluorescence as mock-infected cells. Unexpectedly, eukaryotic cells infected with *E. coli* that did not contain invasin showed the same fluorescent shift as eukaryotic cells infected with *E. coli* containing invasin. These results suggested that either non-specific binding between the *E. coli* and the eukaryotic cells had occurred or that the *E. coli* bind to such cells by a means other than invasin.

Non-specific binding might occur if the eukaryotic cells were not washed properly during the infection. The T-cell line is a suspension cell line, and removal of bacteria requires centrifugation. However, this centrifugation also may pellet the *E. coli* onto the cells and increase non-specific binding. Since the

fibroblast cell line is adherent, most of the non-specific binding should be washed away. The different washing procedures for each cell line may explain our results indicating that 100% of T cells had bacteria attached whereas only 10% of the fibroblast cells had attached bacteria in the absence of invasin. To remove any non-specific binding after infection, the infected eukaryotic cells were centrifuged through a Histopaque gradient, which pellets the unattached *E. coli* but leaves the cells and any attached *E. coli* at the interphase. Comparisons between the Histopaque washing and the previous method of washing indicated that non-specific binding had occurred, yet the majority of the cells retained *E. coli* regardless of invasin expression. These results suggested that *E. coli* binds to eukaryotic cells independent of the invasin. EL4b T cells do not express β1 integrins and, therefore, *E. coli* should not bind these cells if binding requires invasin. Using the Histopaque method, half the EL4b cells still had attached *E. coli*. These results suggest that *E. coli* attach to eukaryotic cells by an undetermined mechanism. Since bacterial delivery needs to target cells specifically, the *E. coli* system must be further modified.

4.2 DEVELOPMENT OF MMTV AS A GENE THERAPY VECTOR

Our original goal was the development of a bacterial vector that would specifically deliver an MMTV-based plasmid carrying a therapeutic gene

specifically to lymphocytes in the gut-associated tissues of mice. This MMTV plasmid then would allow vector amplification in lymphocytes, which would traffic to the mammary epithelium. Therefore, we attempted the development of a minimal plasmid which would accomplish this objective.

4.2.1 GFP-tagged MMTV clones reveal a novel viral function

After MMTV infection of B and T cells in the GALT, these cells traffic to the mammary gland, where dividing mammary epithelial cells are infected during puberty (56,72,142,180). However, little is known about MMTV targeting to the mammary gland. Development of MMTV as a gene therapy vector requires detection of infected cells and their movement. GFP has been utilized in many trafficking studies due to the simplicity of detecting this protein by microscopy. Therefore, GFP was used to tag the MMTV provirus.

Since the MMTV *env* gene is highly expressed, an IRES-GFP cassette was inserted 3' to the *env* gene. This strategy allows both a higher expression level for GFP and expression of an intact Env protein because the IRES enables GFP to be expressed independently from the Env protein. Moreover, the intact Env protein ensures that MMTV replication will occur normally since an Env-GFP fusion protein might unintentionally affect virus-cell interactions. However, the MMTV *env* gene overlaps the *sag* gene, which required duplication of the shared

sequence to allow full-length Sag expression. FACS analysis of different GFP-tagged MMTV clones indicated that GFP was expressed, although at reduced levels compared to the GFP-expression vector. Since GFP expression also was inducible by glucocorticoids, our results suggested that GFP was expressed from the *env* mRNA, which is transcribed from the inducible U3 promoter. Thus, GFP could be used to tag MMTV-infected cells.

4.2.2 GFP-tagged MMTV clones fail to express a *sag* transcript

To determine if insertion of the *IRES-GFP* sequence altered MMTV transcription, RT-PCR analysis of transfected cells was performed. GFP-tagged MMTV clones expressed both *gag* and *env* mRNAs. However, none of the clones produced a *sag* transcript from the intragenic *env* promoter (Fig. 3.16), and each produced an *env-IRES* transcript that was not the expected size. When the shared sequence between *env* and *sag* genes was duplicated, the original splice acceptor for the *sag* gene (SA 8467) was located 5' to the duplicated sequence and the IRES-GFP cassette. To preserve *sag* mRNA splicing, an engineered splice acceptor was inserted upstream of the 3' LTR (Fig. 3.15). Cellular splicing machinery prefers the most 5' splice acceptor and also completely removes all introns from mRNA, consistent with our observations with the GFP-tagged MMTV clones (36,121,153). The unexpected product detected for the *env-IRES*

transcript corresponds to a doubly spliced transcript that uses both splice donor and splice acceptors, and the inability to detect *sag* mRNA most likely resulted from use of the first splice acceptor at position 8467 instead of the engineered second splice acceptor.

The original splice acceptor located 5' to the duplicated sequence was mutated to facilitate correct splicing of the *sag* gene, and FACS analysis showed that GFP expression was maintained. However, RT-PCR analysis again indicated that *sag* mRNA was not expressed despite *gag* and *env* mRNA production. Sequences other than the splice donor and splice acceptor have been shown to be important for recognition by the splicing machinery of the cell, including a polypyrimidine tract and a branch site (Fig. 3.19) (87,153). Lack of such sequences proximal to the engineered splice acceptor may explain failure to produce *sag* transcripts. The probable polypyrimidine tract and branch site for the *sag* mRNA were reinserted 5' to the engineered splice acceptor, yet a *sag* transcript remained undetectable by RT-PCR analysis.

Retroviruses require expression of both spliced and unspliced mRNAs for replication, yet cellular machinery usually splices cellular mRNAs to completion before export to the cytoplasm. To allow expression of both forms of mRNAs, retroviruses must regulate the splicing events that take place in the nucleus (46,134,229). Retroviruses are known to use suboptimal splice sites, including

the polypyrimidine tract and branch site, which ensures that unspliced or incompletely spliced mRNAs are available for export (174,213,224). Wild-type MMTV maintains suboptimal splice sites to produce a *sag* transcript and, therefore, the GFP-tagged MMTV clone that includes the polypyrimidine tract and branch site also should produce *sag* mRNA. However, the actual polypyrimidine tract and branch site have not yet been identified for the MMTV *sag* mRNA, so it is possible that the sequences required for this transcript were inadvertently omitted. Exon sequences also regulate splicing, and both positive and negative elements have been identified in retroviruses (4,5,174,214). The expression of *sag* mRNA from wild-type MMTV suggests that negative splicing elements are not a factor. However, the duplication of the MMTV *env*/*sag* junction in addition to the insertion of the IRES-GFP cassette may have shifted a splicing enhancer element to a location that is too distant to act upon the *sag* splice acceptor, thus preventing correct expression.

4.2.3 GFP-tagged MMTV clones lack Gag protein production due to a defect in viral RNA export

Cells transfected with the GFP-tagged MMTV clones also were tested for the capsid protein and its precursors. Immunoblot analysis indicated that the capsid protein was not expressed since the precursor Pr77 was not detected. These results were surprising since all the changes to the MMTV provirus were

located in the *env* gene, whereas the capsid protein is expressed from the *gag* gene. However, as stated previously, retroviruses require expression of unspliced or partially spliced mRNAs and subsequent export to the cytoplasm for expression of the Gag, Pro, and Pol proteins. This RNA export mechanism differs depending on the complexity of the retrovirus, although elements controlling export are located in the 3' end of the virus, including the *env* gene (45,46,229). To determine if the changes to the *env* gene interfered in the ability to export unspliced transcripts to the cytoplasm, cytoplasmic and nuclear RNA fractions from cells transfected with the GFP-tagged MMTV clones were subjected to RT-PCR analysis. The results showed that the GFP-tagged MMTV clones were unable to export unspliced mRNA to the cytoplasm, although such RNAs were detectable in the nuclear fraction. These experiments suggest that the viral mechanism used to export unspliced mRNA was impeded by the changes in the *env* gene, which led to the inability to express capsid protein.

4.2.4 Transposon mutagenesis of the MMTV provirus

Retroviruses contain *cis*-acting elements such as a packaging and RNA export signals, while complex retroviruses also contain *trans*-acting factors that modulate export (174,186,229). At least one element required for RNA export was altered in the GFP-tagged MMTV provirus. Therefore, identification of such

sequences is necessary for development of an MMTV-based gene therapy vector. Mutagenesis of the infectious MMTV provirus, pHYB-MTV, allows localization of *cis*- or *trans*-acting factors as well as the identification of genes and proteins that are necessary for infection and trafficking to the mammary gland *in vivo.* To determine these requirements, a trimethoprim transposon was inserted randomly into pHYB-MTV, and each MMTV mutant was screened for the location of the insertion by colony PCR. Transposon mutants from different genes were selected and stably transfected into XC cells since previous results have shown that injection of XC cells stably transfected with pHYB-MTV allows productive infection of mice (155,253).

Expression of the stably transfected transposon mutants was verified by RT-PCR and immunoblot analysis before injection into mice. However, some of the transposon mutants did not produce the expected results. The mutant *env* 95 did not express the normal *sag* transcript, and instead a doublet of ca. 600 bp and another band ca. 400 bp was observed. Although the trimethoprim transposon is not located near the splice donor or acceptor for *sag*, it is possible that the insertion disrupted sequences necessary for proper splicing. Alternatively, a splicing enhancer element may have been interrupted or moved farther away to prevent interaction with the splice acceptor, or the insertion of the trimethoprim transposon might have exposed cryptic splice donor acceptor sites that then are

used by the MMTV mutant virus. Gag protein levels of the mutant *env* 95, as well as mutants *env* 12 and *sag* 8, also were considerably bwer than wild type protein levels. Although the insertions were not located in the *gag* gene, it is likely that they are interfering in the export of unspliced mRNA, which is required for expression of Gag proteins.

Stably transfected transposon mutants expressing the expected transcripts and proteins were selected from each gene and injected into BALB/c weanling mice to determine which mutants allowed infection and trafficking to the mammary gland. An MMTV infection can easily be verified by deletion of Sag-reactive T cells (140,155); therefore, the injected mice were tested for deletion of CD4+Vβ14+ T cells, which react with the Sag expressed by pHYB-MTV. No deletion was observed for any of the stably transfected transposon mutants except a mutation in the U3 region of the 5' LTR, which suggested that each gene was necessary for MMTV infection. This result is not surprising for the *gag* and *pol* genes since they express proteins essential for particle formation, replication, and integration. Also since the Sag protein interaction with the CD4+Vβ14+ T cells causes deletion and the *sag* mutants lack a functional Sag, the absence of deletion for these stably transfected transposon mutants was expected. However, the *env* mutant injected into the mice not only interrupted the *env* gene, but also did not

allow high Gag expression levels. Therefore, the lack of deletion observed with *env* mutants could be attributable to a deficiency in Gag or Env protein.

To determine if the injected transposon mutants could traffic to mammary tissue, mammary gland RNA was harvested and analyzed for *gag* and *sag* transcripts. Most of the mutant-injected mice did not produce either mRNA product. However, one injected mouse from each *sag* mutant expressed a *sag* mRNA product. Previous experiments have indicated that Sag is necessary for viral spread in the mammary gland, and our lab has shown that MMTV mutants expressing nonfunctional *sag* genes were able to sporadically infect mammary glands of mice (69,71,90,155,253). The *sag* transposon mutants infected mammary glands approximately 25% of the time, in agreement with previous findings. The ability of *sag* mutants to cause sporadic infections may be due to recombination between the *sag* mutants and endogenous MMTVs of BALB/c mice. Therefore, it is possible that recombination was responsible for deletion of the trimethoprim transposon to allow functional Sag expression and subsequent infection of the mammary gland.

4.2.5 Identification of MMTV sequences necessary for RNA export

Both complex and simple retroviruses require an RNA element in addition to either viral and/or host proteins for the export of unspliced or partially spliced

RNA to the cytoplasm for the subsequent expression of Gag proteins (45,46,229). Previous results indicated that a viral RNA export system might have been disrupted during mutagenesis of the MMTV provirus. RNA export systems are usually found in the 3' end of the virus and insertion of GFP within the 3' half of the MMTV provirus compromised Gag expression. Subsequently, MMTV proviruses containing transposons in the 3' region were utilized to determine specific insertion sites that do not allow Gag expression.

All the mutants examined expressed *gag* mRNA as assessed by RT-PCR from total RNA, yet three of the *env* mutants, *env* 12, *env* 148 and *env* 149, were unable to express Gag protein after transient transfections. The results of transient transfections using mutant *env* 149 did not correspond to the wild-type expression levels observed in stable transfections with the same mutant. It is possible that the transposon in the *env* 149 mutant was inserted adjacent to a sequence necessary for RNA export instead of disrupting the sequence. This insertion might cause a reduction in Gag expression that would be more apparent in transient transfections where only a fraction of the cells receive the provirus. RNA fractionation of cells transfected with each of the three *env* mutants indicated that *gag* mRNA export was decreased when compared to that of wild-type proviruses, suggesting that the loss of Gag expression results from decreased export of the *gag* transcript.

4.2.6 MMTV lacks detectable CTE activity

Viral export systems are different depending on the complexity of the retrovirus (45,46,229). MMTV is classified as a simple retrovirus, and simple retroviruses contain a CTE or CTE-like element, which binds host proteins required for export of the unspliced transcripts (46,83,116,165). Most of these RNA elements are less than 200 bp and are located in untranslated regions near the 3' LTR; however, a larger CTE-like element located in the coding region of the *pol* gene has been identified in a murine endogenous retrovirus (251). MMTV does not contain any untranslated regions near the 3' end of the provirus and, therefore, a potential CTE must be located in a coding region. The three *env* mutants are localized within the MMTV *env* gene, with *env* 12 and *env* 148 located approximately 200 bp apart at the 5' end of *env* and *env* 149 located at the 3' end of *env* directly upstream of the SA for *sag* near the 3' LTR (Figs. 3.24 and 3.37).

MMTV fragments from the 3' end of the provirus were placed in a CTE-dependent luciferase vector to assay for CTE activity. The fragments tested overlapped (i) all three *env* mutant sites, (ii) the two 5' *env* mutant sites, (iii) the one 3' *env* mutant site, or (iv) the 3' MMTV LTR. Such combinations were used to determine if two different elements or a bipartite element were required for RNA export. The CTE-dependent luciferase vectors containing different MMTV

fragments were transiently transfected into XC cells, which express MMTV Gag proteins and must support unspliced RNA export. None of the MMTV fragments rescued luciferase expression as compared to the levels of the intronless-luciferase vector, which expresses luciferase approximately 100-fold less than the intron-containing vector. However, the positive control CTE from MPMV, which should elevate luciferase expression in intronless vectors, increased luciferase levels less than 2-fold in these rat cells. Higher efficiency of the MPMV CTE has been observed in human cell lines (256). Therefore, MMTV fragments were tested for CTE activity in the human 293T cell line, which allows MMTV protein expression. The MPMV CTE increased luciferase levels 4-fold in 293T cells, which is comparable to the levels reported with other CTE-dependent reporter vectors. However, under these conditions, the MMTV fragments did not increase luciferase levels of the intronless vector, suggesting that MMTV lacks a CTE.

4.2.7 Identification of an MMTV *trans*-acting factor for viral RNA export

Complex retroviruses express *trans*-acting proteins that bind to an RNA element on unspliced or partially spliced transcripts for cytoplasmic export. For HIV-1 and HTLV-1, these proteins are Rev and Rex, which bind to the RRE and the RxRE, respectively (3,84,88,174). A similar protein Rec and its binding site, the RcRE, has been identified in the betaretrovirus HERV-K to allow export of

incompletely spliced mRNAs to the cytoplasm (17,18,131,133,134). HERV-K has considerable homology to MMTV (131), suggesting that MMTV might also encode a *trans*-acting protein for unspliced RNA export.

Each of the identified viral export proteins are expressed from a doubly spliced transcript and, therefore, attempts were made to identify such an RNA from MMTV and TBLV-infected cells. TBLV is 98% identical to MMTV but contains a deletion of the negative regulatory elements in the U3 region of the LTR to allow higher expression in T cells (10,32,50). These related viruses should encode the same system for export of unspliced transcripts to the cytoplasm. Novel mRNAs were detected from both XC cells stably transfected with pHYB-MTV as well as Jurkat cells stably transfected with pTBLV. Sequence analysis of these novel mRNAs from each cell line indicated they were identical doubly spliced products using the known splice donors and splice acceptors found in viral RNA. The product of this novel RNA was named Rem.

The predicted protein sequence of the novel *rem* mRNA contained the Env leader peptide, 162 amino acids from gp52 (SU), and 41 amino acids from gp36 (TM). This sequence differs from the other Rev-related proteins since Rem contains the entire leader peptide, including the signal peptidase cleavage site, whereas the Rev-related proteins lack the cleavage site of the leader peptide due to RNA splicing. Rem contains an NLS and NES for possible functions in import

and export in addition to an arginine-rich motif, which is likely to function in RNA binding. Each of these putative domains closely resembles the recognized domains of Rev and the Rev-related proteins, suggesting that MMTV encodes a *trans*-acting protein, Rem, required for unspliced mRNA export to the cytoplasm. However, the domains important for RNA export are all located in the MMTV leader peptide, which usually is cleaved off after insertion of the nascent polypeptide into the ER membrane. If Rem is an MMTV *trans*-acting factor for unspliced RNA export, it will be important to determine the process by which this protein escapes recognition by the signal recognition particle (25,112).

Trans-acting proteins can be identified through complementation analyses (141,216); therefore, a complementation experiment was performed using the two *env* mutants in the *rem* mRNA open reading frame to determine if MMTV encodes a *trans*-acting protein. The addition of wild-type MMTV to XC cells stably transfected with the mutant *env* 148 rescued the export of unspliced transcripts. Moreover, our laboratory recently showed that overexpression of the Rem or Rem-GFP cDNAs in XC cells can increase export of mutant 148 unspliced RNA to the cytoplasm from 7 to 15-fold (J. Mertz, personal communication). These results further indicate that MMTV expresses a *trans*-acting protein, Rem, which functions in export of unspliced mRNA.

Although Rem and the other retroviral proteins involved in RNA export contain similar domains, sequence comparisons showed no significant homology. This is most likely due to the organization of the proteins, which indicates the domains are located in different positions throughout each protein (Fig. 4.1). Rem also is ca. three times larger than the other RNA export proteins, yet the domains involved in RNA export are located only in the N-terminus, or leader peptide. The function of the Rem C-terminus is not presently understood since the other proteins do not contain such a large sequence without any known domains. It is possible that the C-terminus is required for Rem to fold properly and expose the RNA export domains. However, proper folding can only occur once the protein escapes recognition by the signal recognition particle. Alternatively, Rem may be cleaved at the leader peptide to form two functional proteins, one that functions in RNA export and an additional MMTV accessory protein for which the function is unknown.

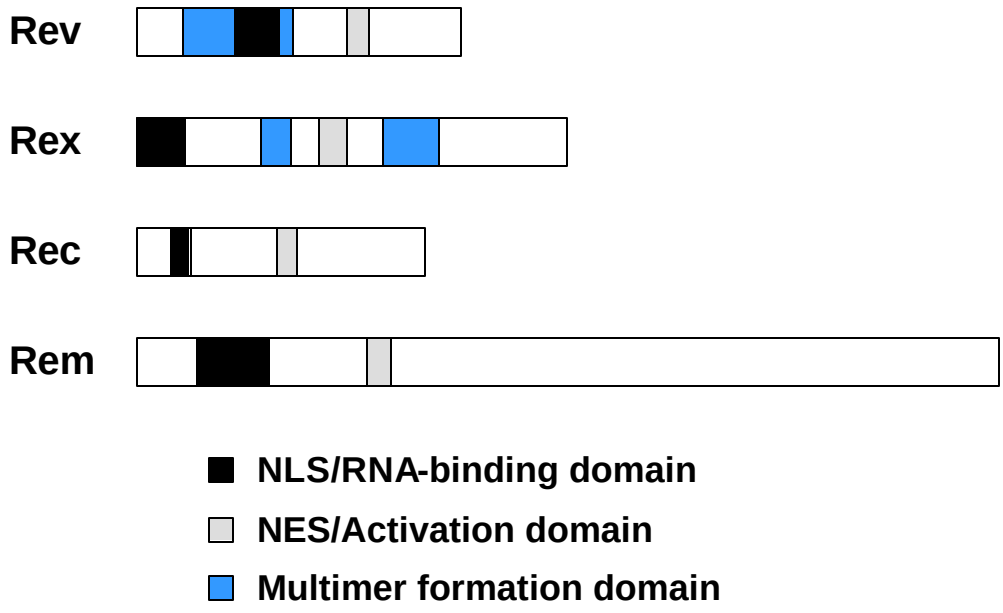


Fig. 4.1. Domain organization of *trans*-acting factors used in RNA transport.

HIV-1 Rev, HTLV-1 Rex, HERV-K Rec, and MMTV Rem are all *trans*-acting factors identified in viral RNA export. Each protein contains an NLS/RNA-binding domain near the N-terminus which functions in nuclear import and binding of the protein to RNA elements on incompletely spliced mRNAs. An NES/Activation domain is located towards the C-terminus, except in Rem, and functions in exporting the protein/RNA complex to the cytoplasm. Rev and Rex also contain multimerization domains important for protein function.

T*rans*-acting proteins involved in RNA export require an RNA element on the unspliced RNA for binding and mRNA export. Such an element would not be identified by complementation analysis. However, previous results have shown that a derivative of the MMTV hybrid provirus (MMTV-neo) was able to be complemented for viral infectivity with wild type MMTV. This MMTV derivative contains only the 3' end of the *env* sequence, starting approximately 40 bp upstream of the 3'LTR, and the 3' LTR(205). The location of RNA elements for HIV-1, HTLV-1, and HERV-K all are in different regions of these retroviruses. The HIV-1 RRE is present only in unspliced or partially spliced mRNAs, specifically in the intron of the *env* sequence, whereas the HTLV-1 RxRE and the HERV-K RcRE are both located in all mRNAs, specifically in the U3R region of the 3' LTR (88,132). From the MMTV derivative data, it seems likely that the MMTV RNA element maps to the 3' LTR, similar to HTLV-1 and HERV-K. The sequence from the MMTV hybrid provirus derivative, other transposon mutants in the 3' LTR, and complementation assays between Rem cDNA and various reporter gene constructs should allow localization of the Rem-responsive element (RmRE).

4.2.8 MMTV is a complex retrovirus

Retroviruses are divided into two categories based on their genome organization, which are (i) simple retroviruses and (ii) complex retroviruses, and both contain coding regions for proteins expressed from *gag*, *pro*, *pol*, and *env*. However, complex retroviruses contain multiply spliced mRNAs that express several accessory proteins, which affect viral RNA synthesis or processing, whereas simple retroviruses contain at most one additional coding region that is not involved in RNA synthesis or processing and is only singly spliced (40).

MMTV is presently classified as a simple retrovirus since it was only known to encode one additional protein, Sag, which is important for MMTV infection and tumorigenesis (191). However, we have now identified a novel, multiply spliced mRNA, *rem*, from MMTV-infected cells, and the Rem predicted protein sequence contains domains similar to RNA export proteins expressed from complex retroviruses (Fig. 4.1). Additionally, MMTV mutants unable to export unspliced mRNA were complemented with both wild-type MMTV and a Rem cDNA plasmid, indicating that MMTV expresses a *trans*-acting protein, Rem, which can overcome an RNA export defect. Since MMTV currently (i) produces a multiply spliced mRNA, (ii) encodes at least two accessory proteins, and (iii) expresses a protein involved in RNA export, it should no longer be

classified as a simple retrovirus and, therefore, MMTV should be reclassified as a complex retrovirus.

4.3 FUTURE DIRECTIONS

MMTV appears to encode a *trans*-acting protein responsible for the export of unspliced mRNA to the cytoplasm, similar to the complex retroviruses HIV-1 and HTLV-1 as well as HERV-K. Much work remains, however, to identify Rem as the protein containing this function. Mutational analysis of the possible NLS and NES from Rem will help determine the import and export capabilities of Rem. Rev and the Rev-like proteins all utilize the Crm1 transport system; therefore it will be important to determine whether MMTV also uses this system or a different RNA or protein transport system. The elucidation of the RNA element utilized by MMTV for export of unspliced transcripts also will help determine important sequences necessary for binding of the Rem protein. These studies will not only allow a better understanding of the replicative ability of MMTV but also will provide locations of both *cis*- and *trans*-acting factors important for the generation of an improved and effective gene therapy vector.

Appendix

List of abbreviations used:

Ab	antibody
ALV	avian leukemia retrovirus
ARE	adenosine/uridine-rich element
ASV	avian sarcoma retrovirus
BGH	bovine growth hormone
BHI	brain-heart infusion
bp	base pairs
C-terminus	carboxy-terminus
$CaCl_2$	calcium chloride
ca.	approximately
CA	capsid protein
CBC	cap-binding complex
CCD	cytochalasin D
CDP	CCAAT displacement protein
CFU	colony-forming units
Crm1	chromosome region maintenance 1
CTE	constitutive transport element
DAP	diaminopimelic acid
DEPC	diethyl pyrocarbonate
Dex	dexamethasone
DHFR	dihydrofolate reductase
DHFR-1	trimethoprim
DMEM	Dulbecco's modified Eagle's medium
DTT	dithiothreitol
E. coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetate
EGTA	ethyleneglycol-bis N,N'-tetra acetic acid
Env	envelope protein
ER	endoplasmic reticulum
FACS	fluorescence-activated cell sorting
FCS	fetal calf serum
FG	phenylalanine-glycine repeats
GALT	gut-associated lymphoid tissue
GFP	green fluorescent protein
HERV-K	human endogenous retrovirus type K
HIV-1	human immunodeficiency virus type 1

hnRNP	heterogeneous nuclear ribonucleoparticle
HRE	hormone response element
HTLV-1	human T-cell leukemia virus type 1
IAP	intracisternal-A particle
IN	integrase protein
kb	kilobases
lacI	*lac* repressor gene
LB	Luria Broth
LSB	low-salt buffer
LTR	long terminal repeat
M cells	membranous epithelial cells
MA	matrix protein
$MgCl_2$	magnesium chloride
MEM	Minimum Essential Medium
MHC	major histocompatibility complex
MHR	major homology region
MLV	murine leukemia virus
MMTV	Mouse mammary tumor virus
MPMV	Mason-Pfizer Monkey virus
mTfR1	mouse transferrin receptor 1
MYM	multiple Yop mutant
N-terminus	amino-terminus
NaCl	sodium chloride
NaOAc	sodium acetate
NaOH	sodium hydroxide
NC	nucleocapsid protein
NES	nuclear export signal
NLS	nuclear localization signal
NPC	nuclear pore complex
NRE	negative regulatory elements
NTF2	nuclear transport factor 2
ONPF	o-Nitrophenyl-b-D-fucoside
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PB	primer binding site
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PMN	polymorphonuclear leukocyte
PMSF	phenylmethylsulfonyl fluoride
PPT	polypurine tract

PR	protease protein
RBD	RNA-binding domain
RcRE	Rec-response element
RIPA	radioimmunoprecipitation assay
Rluc	*Renilla* luciferase
RNP	ribonucleoparticle
RRE	Rev-response element
RT	reverse transcriptase protein
RT-PCR	Reverse-transcriptase PCR
RxRE	Rex-response element
SA	splice acceptor
Sag	Superantigen protein
SATB1	special AT-rich sequence binding protein 1
SD	Splice donor
SDS	sodium dodecyl sulfate
SRV	simian retrovirus
SU	surface protein
TAP	Tip-associated protein
TBLV	Type B leukemogenic virus
TBST	Tris-Buffered Saline Tween-20
TCR	T-cell receptor
TM	transmembrane protein
TTSS	type III secretion system
V_β	β-subunit variable chain
Yop	*Yersinia* outer membrane protein

References

1. **Acha-Orbea, H. and H. R. MacDonald**. 1995. Superantigens of mouse mammary tumor virus. Annu. Rev. Immunol. **13**:459-486.

2. **Acha-Orbea, H. and E. Palmer**. 1991. Mls--a retrovirus exploits the immune system. Immunol. Today **12**:356-361.

3. **Ahmed, Y. F., S. M. Hanly, M. H. Malim, B. R. Cullen, and W. C. Greene**. 1990. Structure-function analyses of the HTLV-I Rex and HIV-1 Rev RNA response elements: insights into the mechanism of Rex and Rev action. Genes Dev. **4**:1014-1022.

4. **Amendt, B. A., D. Hesslein, L. J. Chang, and C. M. Stoltzfus**. 1994. Presence of negative and positive cis-acting RNA splicing elements within and flanking the first tat coding exon of human immunodeficiency virus type 1. Mol. Cell Biol. **14**:3960-3970.

5. **Amendt, B. A., Z. H. Si, and C. M. Stoltzfus**. 1995. Presence of exon splicing silencers within human immunodeficiency virus type 1 tat exon 2 and tat-rev exon 3: evidence for inhibition mediated by cellular factors. Mol. Cell Biol. **15**:4606-4615.

6. **Ardavin, C., F. Luthi, M. Andersson, L. Scarpellino, P. Martin, H. Diggelmann, and H. Acha-Orbea**. 1997. Retrovirus-induced target cell activation in the early phases of infection: the mouse mammary tumor virus model. J. Virol. **71**:7295-7299.

7. **Arts, G. J., S. Kuersten, P. Romby, B. Ehresmann, and I. W. Mattaj**. 1998. The role of exportin-t in selective nuclear export of mature tRNAs. EMBO J. **17**:7430-7441.

8. **Astier, A., H. Avraham, S. N. Manie, J. Groopman, T. Canty, S. Avraham, and A. S. Freedman**. 1997. The related adhesion focal tyrosine kinase is tyrosine-phosphorylated after beta1-integrin stimulation in B cells and binds to p130cas. J. Biol. Chem. **272**:228-232.

9. **Autenrieth, I. B. and R. Firsching**. 1996. Penetration of M cells and destruction of Peyer's patches by *Yersinia enterocolitica*: an ultrastructural and histological study. J. Med. Microbiol. **44**:285-294.

10. **Ball, J. K., L. O. Arthur, and G. A. Dekaban**. 1985. The involvement of a type-B retrovirus in the induction of thymic lymphomas. Virology **140**:159-172.

11. **Bess, J. W., Jr., P. J. Powell, H. J. Issaq, L. J. Schumack, M. K. Grimes, L. E. Henderson, and L. O. Arthur**. 1992. Tightly bound zinc in human immunodeficiency virus type 1, human T-cell leukemia virus type I, and other retroviruses. J. Virol. **66**:840-847.

12. **Beutner, U., E. Kraus, D. Kitamura, K. Rajewsky, and B. T. Huber**. 1994. B cells are essential for murine mammary tumor virus transmission, but not for presentation of endogenous superantigens. J. Exp. Med. **179**:1457-1466.

13. **Bilbao, G., J. Gomez-Navarro, and D. T. Curiel**. 1998. Targeted adenoviral vectors for cancer gene therapy. Adv. Exp. Med. Biol. **451**:365-374.

14. **Bittner, J.J.** 1936. Some possible effects of nursing on the mammary gland tumor incidence in mice. Science **84:**162.

15. **Bittner, J.J.** 1943. Possible relationship of the estrogenic hormones, genetic susceptibility, and milk influence in the production of mammary cancer in mice. Cancer Res. **2**:710-721.

16. **Bliska, J. B., J. E. Galan, and S. Falkow**. 1993. Signal transduction in the mammalian cell during bacterial attachment and entry. Cell **73**:903-920.

17. **Boese, A., U. Galli, M. Geyer, M. Sauter, and N. Mueller-Lantzsch**. 2001. The Rev/Rex homolog HERV-K cORF multimerizes via a C-terminal domain. FEBS Lett. **493**:117-121.

18. **Boese, A., M. Sauter, and N. Mueller-Lantzsch**. 2000. A rev-like NES mediates cytoplasmic localization of HERV-K cORF. FEBS Lett. **468**:65-67.

19. **Boland, A., M. P. Sory, M. Iriarte, C. Kerbourch, P. Wattiau, and G. R. Cornelis**. 1996. Status of YopM and YopN in the *Yersinia* Yop virulon: YopM of *Y.enterocolitica* is internalized inside the cytosol of PU5-1.8 macrophages by the YopB, D, N delivery apparatus. EMBO J. **15**:5191-5201.

20. **Botterman, J. and M. Zabeau**. 1987. A standardized vector system for manipulation and enhanced expression of genes in *Escherichia coli*. DNA **6**:583-591.

21. **Bottone, E. J.** 1997. *Yersinia enterocolitica*: the charisma continues. Clin. Microbiol. Rev. **10**:257-276.

22. **Bouck, J., S. Litwin, A. M. Skalka, and R. A. Katz**. 1998. In vivo selection for intronic splicing signals from a randomized pool. Nucleic Acids Res. **26**:4516-4523.

23. **Boudeau, J., A. L. Glasser, E. Masseret, B. Joly, and A. Darfeuille-Michaud**. 1999. Invasive ability of an *Escherichia coli* strain isolated from the ileal mucosa of a patient with Crohn's disease. Infect. Immun. **67**:4499-4509.

24. **Boulikas, T.** 1993. Nuclear localization signals (NLS). Crit Rev. Eukaryot. Gene Expr. **3**:193-227.

25. **Bovia, F. and K. Strub**. 1996. The signal recognition particle and related small cytoplasmic ribonucleoprotein particles. J. Cell Sci. **109 (Pt 11)**:2601-2608.

26. **Bramblett, D., C. L. Hsu, M. Lozano, K. Earnest, C. Fabritius, and J. Dudley**. 1995. A redundant nuclear protein binding site contributes to negative regulation of the mouse mammary tumor virus long terminal repeat. J. Virol. **69**:7868-7876.

27. **Brandt-Carlson, C. and J. S. Butel**. 1991. Detection and characterization of a glycoprotein encoded by the mouse mammary tumor virus long terminal repeat gene. J. Virol. **65**:6051-6060.

28. **Bray, M., S. Prasad, J. W. Dubay, E. Hunter, K. T. Jeang, D. Rekosh, and M. L. Hammarskjold**. 1994. A small element from the Mason-Pfizer monkey virus genome makes human immunodeficiency virus type 1 expression and replication Rev-independent. Proc. Natl. Acad. Sci. U. S. A **91**:1256-1260.

29. **Brennan, C. M., I. E. Gallouzi, and J. A. Steitz**. 2000. Protein ligands to HuR modulate its interaction with target mRNAs in vivo. J. Cell Biol. **151**:1-14.

30. **Brennan, C. M. and J. A. Steitz**. 2001. HuR and mRNA stability. Cell Mol. Life Sci. **58**:266-277.

31. **Briggs, J. A., B. E. Watson, B. E. Gowen, and S. D. Fuller**. 2004. Cryoelectron microscopy of mouse mammary tumor virus. J. Virol. **78**:2606-2608.

32. **Broussard, D. R., M. M. Lozano, and J. P. Dudley**. 2004. Rorgamma (rorc) is a common integration site in type B leukemogenic virus-induced T-cell lymphomas. J. Virol. **78**:4943-4946.

33. **Bruhn, L., A. Munnerlyn, and R. Grosschedl**. 1997. ALY, a context-dependent coactivator of LEF-1 and AML-1, is required for TCRalpha enhancer function. Genes Dev. **11**:640-653.

34. **Cai, S., H. J. Han, and T. Kohwi-Shigematsu**. 2003. Tissue-specific nuclear architecture and gene expression regulated by SATB1. Nat. Genet. **34**:42-51.

35. **Chaudhuri, R. R., A. M. Khan, and M. J. Pallen**. 2004. coliBASE: an online database for *Escherichia coli*, *Shigella* and *Salmonella* comparative genomics. Nucleic Acids Res. **32 Database issue**:D296-D299.

36. **Chen, S., K. Anderson, and M. J. Moore**. 2000. Evidence for a linear search in bimolecular 3' splice site AG selection. Proc. Natl. Acad. Sci. U. S. A **97**:593-598.

37. **Chinen, J. and J. M. Puck**. 2004. Successes and risks of gene therapy in primary immunodeficiencies. J. Allergy Clin. Immunol. **113**:595-603.

38. **Clouse, K. N., M. J. Luo, Z. Zhou, and R. Reed**. 2001. A Ran-independent pathway for export of spliced mRNA. Nat. Cell Biol. **3**:97-99.

39. **Cochrane, A. W., A. Perkins, and C. A. Rosen**. 1990. Identification of sequences important in the nucleolar localization of human immunodeficiency virus Rev: relevance of nucleolar localization to function. J. Virol. **64**:881-885.

40. **Coffin, J. M., S. H. Hughes, and H. E. Varmus**. 1997. The Interactions of Retroviruses and their Hosts, p. 335-341. *In* J. M. Coffin, S. H. Hughes, and H. E. Varmus (eds.), Retroviruses. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

41. **Cornelis, G. R.** 1998. The *Yersinia* deadly kiss. J. Bacteriol. **180**:5495-5504.

42. **Cornelis, G. R., A. Boland, A. P. Boyd, C. Geuijen, M. Iriarte, C. Neyt, M. P. Sory, and I. Stainier**. 1998. The virulence plasmid of *Yersinia*, an antihost genome. Microbiol. Mol. Biol. Rev. **62**:1315-1352.

43. **Courvalin, P., S. Goussard, and C. Grillot-Courvalin**. 1995. Gene transfer from bacteria to mammalian cells. C. R. Acad. Sci. III **318**:1207-1212.

44. **Cover, T. L. and R. C. Aber**. 1989. *Yersinia enterocolitica*. N. Engl. J. Med. **321**:16-24.

45. **Cullen, B. R.** 1998. Retroviruses as model systems for the study of nuclear RNA export pathways. Virology **249**:203-210.

46. **Cullen, B. R.** 2003. Nuclear mRNA export: insights from virology. Trends Biochem. Sci. **28**:419-424.

47. **Dagert, M. and S. D. Ehrlich**. 1979. Prolonged incubation in calcium chloride improves the competence of *Escherichia coli* cells. Gene **6**:23-28.

48. **Daly, T. J., R. C. Doten, P. Rennert, M. Auer, H. Jaksche, A. Donner, G. Fisk, and J. R. Rusche**. 1993. Biochemical characterization of binding of multiple HIV-1 Rev monomeric proteins to the Rev responsive element. Biochemistry **32**:10497-10505.

49. **Darlix, J. L., C. Gabus, and B. Allain**. 1992. Analytical study of avian reticuloendotheliosis virus dimeric RNA generated in vivo and in vitro. J. Virol. **66**:7245-7252.

50. **Dekaban, G. A. and J. K. Ball**. 1984. Integration of type B retroviral DNA in virus-induced primary murine thymic lymphomas. J. Virol. **52**:784-792.

51. **Denis, F., N. H. Shoukry, M. Delcourt, J. Thibodeau, N. Labrecque, H. McGrath, J. S. Munzer, N. G. Seidah, and R. P. Sekaly**. 2000. Alternative proteolytic processing of mouse mammary tumor virus superantigens. J. Virol. **74**:3067-3073.

52. **Dickson, C. and M. Atterwill**. 1978. Polyproteins related to the major core protein of mouse mammary tumor virus. J. Virol. **26**:660-672.

53. **Dickson, C. and M. Atterwill**. 1980. Structure and processing of the mouse mammary tumor virus glycoprotein precursor pr73env. J. Virol. **35**:349-361.

54. **Donehower, L. A., A. L. Huang, and G. L. Hager**. 1981. Regulatory and coding potential of the mouse mammary tumor virus long terminal redundancy. J. Virol. **37**:226-238.

55. **Dorfman, T., F. Mammano, W. A. Haseltine, and H. G. Gottlinger**. 1994. Role of the matrix protein in the virion association of the human immunodeficiency virus type 1 envelope glycoprotein. J. Virol. **68**:1689-1696.

56. **Dzuris, J. L., T. V. Golovkina, and S. R. Ross**. 1997. Both T and B cells shed infectious mouse mammary tumor virus. J. Virol. **71**:6044-6048.

57. **Eckert, D. M. and P. S. Kim**. 2001. Mechanisms of viral membrane fusion and its inhibition. Annu. Rev. Biochem. **70**:777-810.

58. **Einfeld, D. A. and E. Hunter**. 1994. Expression of the TM protein of Rous sarcoma virus in the absence of SU shows that this domain is capable of oligomerization and intracellular transport. J. Virol. **68**:2513-2520.

59. **Elliott, J. F., B. Pohajdak, D. J. Talbot, J. Shaw, and V. Paetkau**. 1988. Phorbol diester-inducible, cyclosporine-suppressible transcription from a novel promoter within the mouse mammary tumor virus env gene. J. Virol. **62**:1373-1380.

60. **Ernst, R. K., M. Bray, D. Rekosh, and M. L. Hammarskjold**. 1997. A structured retroviral RNA element that mediates nucleocytoplasmic export of intron-containing RNA. Mol. Cell Biol. **17**:135-144.

61. **Ernst, R. K., M. Bray, D. Rekosh, and M. L. Hammarskjold**. 1997. Secondary structure and mutational analysis of the Mason-Pfizer monkey virus RNA constitutive transport element. RNA. **3**:210-222.

62. **Fabre, E. and E. Hurt**. 1997. Yeast genetics to dissect the nuclear pore complex and nucleocytoplasmic trafficking. Annu. Rev. Genet. **31**:277-313.

63. **Fischer, U., J. Huber, W. C. Boelens, I. W. Mattaj, and R. Luhrmann**. 1995. The HIV-1 Rev activation domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs. Cell **82**:475-483.

64. **Flint, S. J., L. W. Enquist, V. R. Racaniello, and A. M. Skalka**. 2004. Reverse Transcription and Integration, p. 217-250. *In* S. J. Flint, L. W. Enquist, V. R. Racaniello, and A. M. Skalka (eds.), Principles of Virology. ASM Press, Washington, DC.

65. **Fornerod, M., M. Ohno, M. Yoshida, and I. W. Mattaj**. 1997. CRM1 is an export receptor for leucine-rich nuclear export signals. Cell **90**:1051-1060.

66. **Forsberg, A., R. Rosqvist, and H. Wolf-Watz**. 1994. Regulation and polarized transfer of the *Yersinia* outer proteins (Yops) involved in antiphagocytosis. Trends Microbiol. **2**:14-19.

67. **Gallouzi, I. E., C. M. Brennan, and J. A. Steitz**. 2001. Protein ligands mediate the CRM1-dependent export of HuR in response to heat shock. RNA. **7**:1348-1361.

68. **Gallouzi, I. E. and J. A. Steitz**. 2001. Delineation of mRNA export pathways by the use of cell-permeable peptides. Science **294**:1895-1901.

69. **Golovkina, T. V., A. Chervonsky, J. P. Dudley, and S. R. Ross**. 1992. Transgenic mouse mammary tumor virus superantigen expression prevents viral infection. Cell **69**:637-645.

70. **Golovkina, T. V., A. Chervonsky, J. A. Prescott, C. A. Janeway, Jr., and S. R. Ross**. 1994. The mouse mammary tumor virus envelope gene product is required for superantigen presentation to T cells. J. Exp. Med. **179**:439-446.

71. **Golovkina, T. V., J. P. Dudley, A. B. Jaffe, and S. R. Ross**. 1995. Mouse mammary tumor viruses with functional superantigen genes are selected during in vivo infection. Proc. Natl. Acad. Sci. U. S. A **92**:4828-4832.

72. **Golovkina, T. V., J. P. Dudley, and S. R. Ross**. 1998. B and T cells are required for mouse mammary tumor virus spread within the mammary gland. J. Immunol. **161**:2375-2382.

73. **Golovkina, T. V., J. Dzuris, H. B. van den, A. B. Jaffe, P. C. Wright, S. M. Cofer, and S. R. Ross**. 1998. A novel membrane protein is a mouse mammary tumor virus receptor. J. Virol. **72**:3066-3071.

74. **Gorlich, D.** 1998. Transport into and out of the cell nucleus. EMBO J. **17**:2721-2727.

75. **Gorlich, D. and U. Kutay**. 1999. Transport between the cell nucleus and the cytoplasm. Annu. Rev. Cell Dev. Biol. **15**:607-660.

76. **Gorlich, D., N. Pante, U. Kutay, U. Aebi, and F. R. Bischoff**. 1996. Identification of different roles for RanGDP and RanGTP in nuclear protein import. EMBO J. **15**:5584-5594.

77. **Grimm, S. L. and S. K. Nordeen**. 1998. Mouse mammary tumor virus sequences responsible for activating cellular oncogenes. J. Virol. **72**:9428-9435.

78. **Gruter, P., C. Tabernero, C. von Kobbe, C. Schmitt, C. Saavedra, A. Bachi, M. Wilm, B. K. Felber, and E. Izaurralde**. 1998. TAP, the human homolog of Mex67p, mediates CTE-dependent RNA export from the nucleus. Mol. Cell **1**:649-659.

79. **Gunzburg, W. H., F. Heinemann, S. Wintersperger, T. Miethke, H. Wagner, V. Erfle, and B. Salmons**. 1993. Endogenous superantigen expression controlled by a novel promoter in the MMTV long terminal repeat. Nature **364**:154-158.

80. **Hacein-Bey-Abina, S., C. von Kalle, M. Schmidt, M. P. McCormack, N. Wulffraat, P. Leboulch, A. Lim, C. S. Osborne, R. Pawliuk, E. Morillon, R. Sorensen, A. Forster, P. Fraser, J. I. Cohen, B. G. de Saint, I. Alexander, U. Wintergerst, T. Frebourg, A. Aurias, D. Stoppa-Lyonnet, S. Romana, I. Radford-Weiss, F. Gross, F. Valensi, E. Delabesse, E. Macintyre, F. Sigaux, J. Soulier, L. E. Leiva, M. Wissler, C. Prinz, T. H. Rabbitts, F. Le Deist, A. Fischer, and M. Cavazzana-Calvo**. 2003. LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. Science **302**:415-419.

81. **Hadzopoulou-Cladaras, M., B. K. Felber, C. Cladaras, A. Athanassopoulos, A. Tse, and G. N. Pavlakis**. 1989. The rev (trs/art) protein of human immunodeficiency virus type 1 affects viral mRNA and protein expression via a cis-acting sequence in the env region. J. Virol. **63**:1265-1274.

82. **Hakansson, S., E. E. Galyov, R. Rosqvist, and H. Wolf-Watz**. 1996. The *Yersinia* YpkA Ser/Thr kinase is translocated and subsequently targeted to

the inner surface of the HeLa cell plasma membrane. Mol. Microbiol. **20**:593-603.

83. **Hammarskjold, M. L.** 2001. Constitutive transport element-mediated nuclear export. Curr. Top. Microbiol. Immunol. **259**:77-93.

84. **Hanly, S. M., L. T. Rimsky, M. H. Malim, J. H. Kim, J. Hauber, D. M. Duc, S. Y. Le, J. V. Maizel, B. R. Cullen, and W. C. Greene**. 1989. Comparative analysis of the HTLV-I Rex and HIV-1 Rev trans-regulatory proteins and their RNA response elements. Genes Dev. **3**:1534-1544.

85. **Hanson, L. A.** 1976. *Esch. coli* infections in childhood. Significance of bacterial virulence and immune defence. Arch. Dis. Child **51**:737-742.

86. **Hartig, E., B. Nierlich, S. Mink, G. Nebl, and A. C. Cato**. 1993. Regulation of expression of mouse mammary tumor virus through sequences located in the hormone response element: involvement of cell-cell contact and a negative regulatory factor. J. Virol. **67**:813-821.

87. **Hastings, M. L. and A. R. Krainer**. 2001. Pre-mRNA splicing in the new millennium. Curr. Opin. Cell Biol. **13**:302-309.

88. **Hauber, J.** 2001. Nuclear export mediated by the Rev/Rex class of retroviral Trans-activator proteins. Curr. Top. Microbiol. Immunol. **259**:55-76.

89. **Heaphy, S., C. Dingwall, I. Ernberg, M. J. Gait, S. M. Green, J. Karn, A. D. Lowe, M. Singh, and M. A. Skinner**. 1990. HIV-1 regulator of virion expression (Rev) protein binds to an RNA stem-loop structure located within the Rev response element region. Cell **60**:685-693.

90. **Held, W., G. A. Waanders, A. N. Shakhov, L. Scarpellino, H. Acha-Orbea, and H. R. MacDonald**. 1993. Superantigen-induced immune stimulation amplifies mouse mammary tumor virus infection and allows virus transmission. Cell **74**:529-540.

91. **Henderson, B. R. and P. Percipalle**. 1997. Interactions between HIV Rev and nuclear import and export factors: the Rev nuclear localisation signal mediates specific binding to human importin-beta. J. Mol. Biol. **274**:693-707.

92. **Henderson, L. E., R. Sowder, G. Smythers, and S. Oroszlan**. 1983. Terminal amino acid sequences and proteolytic cleavage sites of mouse mammary tumor virus env gene products. J. Virol. **48**:314-319.

93. **Hizi, A., L. E. Henderson, T. D. Copeland, R. C. Sowder, H. C. Krutzsch, and S. Oroszlan**. 1989. Analysis of gag proteins from mouse mammary tumor virus. J. Virol. **63**:2543-2549.

94. **Hope, T. J.** 1999. The ins and outs of HIV Rev. Arch. Biochem. Biophys. **365**:186-191.

95. **Hsu, P. N., B. P. Wolf, N. Sutkowski, B. McLellan, H. L. Ploegh, and B. T. Huber**. 2001. Association of mouse mammary tumor virus superantigen with MHC class II during biosynthesis. J. Immunol. **166**:3309-3314.

96. **Hu, W. S. and V. K. Pathak**. 2000. Design of retroviral vectors and helper cells for gene therapy. Pharmacol. Rev. **52**:493-511.

97. **Hurt, E., S. Hannus, B. Schmelzl, D. Lau, D. Tollervey, and G. Simos**. 1999. A novel in vivo assay reveals inhibition of ribosomal nuclear export in ran-cycle and nucleoporin mutants. J. Cell Biol. **144**:389-401.

98. **Iriarte, M. and G. R. Cornelis**. 1998. YopT, a new *Yersinia* Yop effector protein, affects the cytoskeleton of host cells. Mol. Microbiol. **29**:915-929.

99. **Isberg, R. R.** 1990. Pathways for the penetration of enteroinvasive *Yersinia* into mammalian cells. Mol. Biol. Med. **7**:73-82.

100. **Isberg, R. R. and S. Falkow**. 1985. A single genetic locus encoded by *Yersinia pseudotuberculosis* permits invasion of cultured animal cells by *Escherichia coli* K-12. Nature **317**:262-264.

101. **Isberg, R. R. and J. M. Leong**. 1990. Multiple beta 1 chain integrins are receptors for invasin, a protein that promotes bacterial penetration into mammalian cells. Cell **60**:861-871.

102. **Izaurralde, E. and S. Adam**. 1998. Transport of macromolecules between the nucleus and the cytoplasm. RNA. **4**:351-364.

103. **Izaurralde, E., A. Jarmolowski, C. Beisel, I. W. Mattaj, G. Dreyfuss, and U. Fischer**. 1997. A role for the M9 transport signal of hnRNP A1 in mRNA nuclear export. J. Cell Biol. **137**:27-35.

104. **Izaurralde, E., J. Lewis, C. Gamberi, A. Jarmolowski, C. McGuigan, and I. W. Mattaj**. 1995. A cap-binding protein complex mediating U snRNA export. Nature **376**:709-712.

105. **Jacks, T., K. Townsley, H. E. Varmus, and J. Majors**. 1987. Two efficient ribosomal frameshifting events are required for synthesis of mouse mammary tumor virus gag-related polyproteins. Proc. Natl. Acad. Sci. U.S.A **84**:4298-4302.

106. **Jackson, B. and C. C. Little**. 1933. The existence of non-chromosomal influence in the incidence of mammary tumors in mice. Science **78:**465-466.

107. **Jarmolowski, A., W. C. Boelens, E. Izaurralde, and I. W. Mattaj**. 1994. Nuclear export of different classes of RNA is mediated by specific factors. J. Cell Biol. **124**:627-635.

108. **Joshi, S., A. Van Brunschot, I. Robson, and A. Bernstein**. 1990. Efficient replication, integration, and packaging of retroviral vectors with modified long terminal repeats containing the packaging signal. Nucleic Acids Res. **18**:4223-4226.

109. **Kadowaki, T., M. Hitomi, S. Chen, and A. M. Tartakoff**. 1994. Nuclear mRNA accumulation causes nucleolar fragmentation in yeast mtr2 mutant. Mol. Biol. Cell **5**:1253-1263.

110. **Kang, Y. and B. R. Cullen**. 1999. The human Tap protein is a nuclear mRNA export factor that contains novel RNA-binding and nucleocytoplasmic transport sequences. Genes Dev. **13**:1126-1139.

111. **Katoh, M.** 2002. WNT and FGF gene clusters (review). Int. J. Oncol. **21**:1269-1273.

112. **Keenan, R. J., D. M. Freymann, R. M. Stroud, and P. Walter**. 2001. The signal recognition particle. Annu. Rev. Biochem. **70**:755-775.

113. **Kim, F. J., A. A. Beeche, J. J. Hunter, D. J. Chin, and T. J. Hope**. 1996. Characterization of the nuclear export signal of human T-cell lymphotropic virus type 1 Rex reveals that nuclear export is mediated by position-variable hydrophobic interactions. Mol. Cell Biol. **16**:5147-5155.

114. **Kim, S. H., S. Kim, and P. D. Robbins**. 2000. Retroviral vectors. Adv. Virus Res. **55**:545-563.

115. **Klemenz, R., M. Reinhardt, and H. Diggelmann**. 1981. Sequence determination of the 3' end of mouse mammary tumor virus RNA. Mol. Biol. Rep. **7**:123-126.

116. **Kootstra, N. A. and I. M. Verma**. 2003. Gene therapy with viral vectors. Annu. Rev. Pharmacol. Toxicol. **43**:413-439.

117. **Kunkel, E. J., D. J. Campbell, and E. C. Butcher**. 2003. Chemokines in lymphocyte trafficking and intestinal immunity. Microcirculation. **10**:313-323.

118. **Kutay, U., F. R. Bischoff, S. Kostka, R. Kraft, and D. Gorlich**. 1997. Export of importin alpha from the nucleus is mediated by a specific nuclear transport factor. Cell **90**:1061-1071.

119. **Kutay, U., G. Lipowsky, E. Izaurralde, F. R. Bischoff, P. Schwarzmaier, E. Hartmann, and D. Gorlich**. 1998. Identification of a tRNA-specific nuclear export receptor. Mol. Cell **1**:359-369.

120. **LaFemina, R. L., P. L. Callahan, and M. G. Cordingley**. 1991. Substrate specificity of recombinant human immunodeficiency virus integrase protein. J. Virol. **65**:5624-5630.

121. **Le Hir, H., A. Nott, and M. J. Moore**. 2003. How introns influence and enhance eukaryotic gene expression. Trends Biochem. Sci. **28**:215-220.

122. **Ledley, F. D.** 1995. Nonviral gene therapy: the promise of genes as pharmaceutical products. Hum. Gene Ther. **6**:1129-1144.

123. **Lee, E. G., A. Alidina, C. May, and M. L. Linial**. 2003. Importance of basic residues in binding of rous sarcoma virus nucleocapsid to the RNA packaging signal. J. Virol. **77**:2010-2020.

124. **Lee, V. T. and O. Schneewind**. 1999. Type III secretion machines and the pathogenesis of enteric infections caused by *Yersinia* and *Salmonella* spp. Immunol. Rev. **168**:241-255.

125. **Lei, E. P. and P. A. Silver**. 2002. Protein and RNA export from the nucleus. Dev. Cell **2**:261-272.

126. **Li, J., H. Tang, T. M. Mullen, C. Westberg, T. R. Reddy, D. W. Rose, and F. Wong-Staal**. 1999. A role for RNA helicase A in post-

transcriptional regulation of HIV type 1. Proc. Natl. Acad. Sci. U. S. A **96**:709-714.

127. **Li, S. and L. Huang**. 2000. Nonviral gene therapy: promises and challenges. Gene Ther. **7**:31-34.

128. **Liu, J., D. Bramblett, Q. Zhu, M. Lozano, R. Kobayashi, S. R. Ross, and J. P. Dudley**. 1997. The matrix attachment region-binding protein SATB1 participates in negative regulation of tissue-specific gene expression. Mol. Cell Biol. **17**:5275-5287.

129. **Liu, S. L., F. M. Duh, M. I. Lerman, and A. D. Miller**. 2003. Role of virus receptor Hyal2 in oncogenic transformation of rodent fibroblasts by sheep betaretrovirus env proteins. J. Virol. **77**:2850-2858.

130. **Lower, R., K. Boller, B. Hasenmaier, C. Korbmacher, N. Muller-Lantzsch, J. Lower, and R. Kurth**. 1993. Identification of human endogenous retroviruses with complex mRNA expression and particle formation. Proc. Natl. Acad. Sci. U. S. A **90**:4480-4484.

131. **Lower, R., R. R. Tonjes, C. Korbmacher, R. Kurth, and J. Lower**. 1995. Identification of a Rev-related protein by analysis of spliced transcripts of the human endogenous retroviruses HTDV/HERV-K. J. Virol. **69**:141-149.

132. **Magin-Lachmann, C., S. Hahn, H. Strobel, U. Held, J. Lower, and R. Lower**. 2001. Rec (formerly Corf) function requires interaction with a complex, folded RNA structure within its responsive element rather than binding to a discrete specific binding site. J. Virol. **75**:10359-10371.

133. **Magin, C., J. Hesse, J. Lower, and R. Lower**. 2000. Corf, the Rev/Rex homologue of HTDV/HERV-K, encodes an arginine-rich nuclear localization signal that exerts a trans-dominant phenotype when mutated. Virology **274**:11-16.

134. **Magin, C., R. Lower, and J. Lower**. 1999. cORF and RcRE, the Rev/Rex and RRE/RxRE homologues of the human endogenous retrovirus family HTDV/HERV-K. J. Virol. **73**:9496-9507.

135. **Majors, J. and H. E. Varmus**. 1983. A small region of the mouse mammary tumor virus long terminal repeat confers glucocorticoid hormone regulation on a linked heterologous gene. Proc. Natl. Acad. Sci. U. S. A **80**:5866-5870.

136. **Majors, J. E. and H. E. Varmus**. 1983. Nucleotide sequencing of an apparent proviral copy of env mRNA defines determinants of expression of the mouse mammary tumor virus env gene. J. Virol. **47**:495-504.

137. **Malim, M. H. and B. R. Cullen**. 1991. HIV-1 structural gene expression requires the binding of multiple Rev monomers to the viral RRE: implications for HIV-1 latency. Cell **65**:241-248.

138. **Malim, M. H., J. Hauber, S. Y. Le, J. V. Maizel, and B. R. Cullen**. 1989. The HIV-1 rev trans-activator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA. Nature **338**:254-257.

139. **Mantle, M., L. Basaraba, S. C. Peacock, and D. G. Gall**. 1989. Binding of *Yersinia enterocolitica* to rabbit intestinal brush border membranes, mucus, and mucin. Infect. Immun. **57**:3292-3299.

140. **Marrack, P., E. Kushnir, and J. Kappler**. 1991. A maternally inherited superantigen encoded by a mammary tumour virus. Nature **349**:524-526.

141. **Martarano, L., R. Stephens, N. Rice, and D. Derse**. 1994. Equine infectious anemia virus trans-regulatory protein Rev controls viral mRNA stability, accumulation, and alternative splicing. J. Virol. **68**:3102-3111.

142. **Matsuzawa, A., H. Nakano, T. Yoshimoto, and K. Sayama**. 1995. Biology of mouse mammary tumor virus (MMTV). Cancer Lett. **90**:3-11.

143. **Mattaj, I. W. and L. Englmeier**. 1998. Nucleocytoplasmic transport: the soluble phase. Annu. Rev. Biochem. **67**:265-306.

144. **May, F. E. and B. R. Westley**. 1986. Structure of a human retroviral sequence related to mouse mammary tumor virus. J. Virol. **60**:743-749.

145. **Menendez-Arias, L., C. Risco, d. S. Pinto, and S. Oroszlan**. 1992. Purification of immature cores of mouse mammary tumor virus and immunolocalization of protein domains. J. Virol. **66**:5615-5620.

146. **Mermod, J. J., S. Bourgeois, N. Defer, and M. Crepin**. 1983. Demethylation and expression of murine mammary tumor proviruses in mouse thymoma cell lines. Proc. Natl. Acad. Sci. U. S. A **80**:110-114.

147. **Michael, W. M.** 2000. Nucleocytoplasmic shuttling signals: two for the price of one. Trends Cell Biol. **10**:46-50.

148. **Michalides, R., A. van Ooyen, and R. Nusse**. 1983. Mouse mammary tumor virus expression and mammary tumor development. Curr. Top. Microbiol. Immunol. **106**:57-78.

149. **Mielke, M. E., S. Ehlers, and H. Hahn**. 1993. The role of cytokines in experimental listeriosis. Immunobiology **189**:285-315.

150. **Mohan, N., D. Mottershead, M. Subramanyam, U. Beutner, and B. T. Huber**. 1993. Production and characterization of an Mls-1-specific monoclonal antibody. J. Exp. Med. **177**:351-358.

151. **Mok, E., T. V. Golovkina, and S. R. Ross**. 1992. A mouse mammary tumor virus mammary gland enhancer confers tissue-specific but not lactation-dependent expression in transgenic mice. J. Virol. **66**:7529-7532.

152. **Molina, R. M., Y. Chebloune, G. Verdier, and C. Legras**. 1995. Influence of expression and cis-acting sequences from avian leukosis viruses (ALVs) on stability of (ALV)-based retrovirus vectors. C. R. Acad. Sci. III **318**:541-551.

153. **Moore, M. J.** 2000. Intron recognition comes of AGe. Nat. Struct. Biol. **7**:14-16.

154. **Moy, T. I. and P. A. Silver**. 1999. Nuclear export of the small ribosomal subunit requires the ran-GTPase cycle and certain nucleoporins. Genes Dev. **13**:2118-2133.

155. **Mustafa, F., M. Lozano, and J. P. Dudley**. 2000. C3H mouse mammary tumor virus superantigen function requires a splice donor site in the envelope gene. J. Virol. **74**:9431-9440.

156. **Nakielny, S. and G. Dreyfuss**. 1999. Transport of proteins and RNAs in and out of the nucleus. Cell **99**:677-690.

157. **Nakielny, S., U. Fischer, W. M. Michael, and G. Dreyfuss**. 1997. RNA transport. Annu. Rev. Neurosci. **20**:269-301.

158. **Nhieu, G. T. and P. J. Sansonetti**. 1999. Mechanism of Shigella entry into epithelial cells. Curr. Opin. Microbiol. **2**:51-55.

159. **Niidome, T. and L. Huang**. 2002. Gene therapy progress and prospects: nonviral vectors. Gene Ther. **9**:1647-1652.

160. **Noriega, F. R., G. Losonsky, C. Lauderbaugh, F. M. Liao, J. Y. Wang, and M. M. Levine**. 1996. Engineered deltaguaB-A deltavirG *Shigella flexneri* 2a strain CVD 1205: construction, safety, immunogenicity, and potential efficacy as a mucosal vaccine. Infect. Immun. **64**:3055-3061.

161. **Noriega, F. R., J. Y. Wang, G. Losonsky, D. R. Maneval, D. M. Hone, and M. M. Levine**. 1994. Construction and characterization of attenuated delta aroA delta virG *Shigella flexneri* 2a strain CVD 1203, a prototype live oral vaccine. Infect. Immun. **62**:5168-5172.

162. **Ogert, R. A. and K. L. Beemon**. 1998. Mutational analysis of the rous sarcoma virus DR posttranscriptional control element. J. Virol. **72**:3407-3411.

163. **Ogert, R. A., L. H. Lee, and K. L. Beemon**. 1996. Avian retroviral RNA element promotes unspliced RNA accumulation in the cytoplasm. J. Virol. **70**:3834-3843.

164. **Paca, R. E., R. A. Ogert, C. S. Hibbert, E. Izaurralde, and K. L. Beemon**. 2000. Rous sarcoma virus DR posttranscriptional elements use a novel RNA export pathway. J. Virol. **74**:9507-9514.

165. **Palu, G., C. Parolin, Y. Takeuchi, and M. Pizzato**. 2000. Progress with retroviral gene vectors. Rev. Med. Virol. **10**:185-202.

166. **Pasquinelli, A. E., R. K. Ernst, E. Lund, C. Grimm, M. L. Zapp, D. Rekosh, M. L. Hammarskjold, and J. E. Dahlberg**. 1997. The constitutive transport element (CTE) of Mason-Pfizer monkey virus (MPMV) accesses a cellular mRNA export pathway. EMBO J. **16**:7500-7510.

167. **Pearl, L. H. and W. R. Taylor**. 1987. Sequence specificity of retroviral proteases. Nature **328**:482.

168. **Pepinsky, R. B. and V. M. Vogt**. 1984. Fine-structure analyses of lipid-protein and protein-protein interactions of gag protein p19 of the avian sarcoma and leukemia viruses by cyanogen bromide mapping. J. Virol. **52**:145-153.

169. **Peters, G. and C. Glover**. 1980. tRNA's and priming of RNA-directed DNA synthesis in mouse mammary tumor virus. J. Virol. **35**:31-40.

170. **Pfeifer, A. and I. M. Verma**. 2001. Gene therapy: promises and problems. Annu. Rev. Genomics Hum. Genet. **2**:177-211.

171. **Pilarski, L. M., B. R. Yacyshyn, G. S. Jensen, E. Pruski, and H. F. Pabst**. 1991. Beta 1 integrin (CD29) expression on human postnatal T cell subsets defined by selective CD45 isoform expression. J. Immunol. **147**:830-837.

172. **Pinol-Roma, S. and G. Dreyfuss**. 1991. Transcription-dependent and transcription-independent nuclear transport of hnRNP proteins. Science **253**:312-314.

173. **Pinol-Roma, S. and G. Dreyfuss**. 1992. Shuttling of pre-mRNA binding proteins between nucleus and cytoplasm. Nature **355**:730-732.

174. **Pollard, V. W. and M. H. Malim**. 1998. The HIV-1 Rev protein. Annu. Rev. Microbiol. **52**:491-532.

175. **Popa, I., M. E. Harris, J. E. Donello, and T. J. Hope**. 2002. CRM1-dependent function of a cis-acting RNA export element. Mol. Cell Biol. **22**:2057-2067.

176. **Pornillos, O., J. E. Garrus, and W. I. Sundquist**. 2002. Mechanisms of enveloped RNA virus budding. Trends Cell Biol. **12**:569-579.

177. **Portman, D. S., J. P. O'Connor, and G. Dreyfuss**. 1997. YRA1, an essential Saccharomyces cerevisiae gene, encodes a novel nuclear protein with RNA annealing activity. RNA. **3**:527-537.

178. **Pringle, C. R.** 1999. Virus taxonomy--1999. The universal system of virus taxonomy, updated to include the new proposals ratified by the International Committee on Taxonomy of Viruses during 1998. Arch. Virol. **144**:421-429.

179. **Pupo, G. M., D. K. Karaolis, R. Lan, and P. R. Reeves**. 1997. Evolutionary relationships among pathogenic and nonpathogenic Escherichia coli strains inferred from multilocus enzyme electrophoresis and mdh sequence studies. Infect. Immun. **65**:2685-2692.

180. **Qin, W., T. V. Golovkina, T. Peng, I. Nepomnaschy, V. Buggiano, I. Piazzon, and S. R. Ross**. 1999. Mammary gland expression of mouse mammary tumor virus is regulated by a novel element in the long terminal repeat. J. Virol. **73**:368-376.

181. **Racevskis, J. and O. Prakash**. 1984. Proteins encoded by the long terminal repeat region of mouse mammary tumor virus: identification by hybrid-selected translation. J. Virol. **51**:604-610.

182. **Ragheb, J. A. and W. F. Anderson**. 1994. Uncoupled expression of Moloney murine leukemia virus envelope polypeptides SU and TM: a functional analysis of the role of TM domains in viral entry. J. Virol. **68**:3207-3219.

183. **Rattray, A. J. and J. J. Champoux**. 1987. The role of Moloney murine leukemia virus RNase H activity in the formation of plus-strand primers. J. Virol. **61**:2843-2851.

184. **Reed, R. and E. Hurt**. 2002. A conserved mRNA export machinery coupled to pre-mRNA splicing. Cell **108**:523-531.

185. **Reilly, M., D. Mix, and G. M. Winslow**. 2000. Detection of viral superantigen-class II MHC interactions at the cell surface. Mol. Immunol. **37**:987-993.

186. **Rein, A.** 1994. Retroviral RNA packaging: a review. Arch. Virol. Suppl **9**:513-522.

187. **Resnick, R., C. A. Omer, and A. J. Faras**. 1984. Involvement of retrovirus reverse transcriptase-associated RNase H in the initiation of strong-stop (+) DNA synthesis and the generation of the long terminal repeat. J. Virol. **51**:813-821.

188. **Robbins, J., B. J. Blondel, D. Gallahan, and R. Callahan**. 1992. Mouse mammary tumor gene int-3: a member of the notch gene family transforms mammary epithelial cells. J. Virol. **66**:2594-2599.

189. **Robbins, P. D. and S. C. Ghivizzani**. 1998. Viral vectors for gene therapy. Pharmacol. Ther. **80**:35-47.

190. **Rodrigues, J. P., M. Rode, D. Gatfield, B. J. Blencowe, M. Carmo-Fonseca, and E. Izaurralde**. 2001. REF proteins mediate the export of spliced and unspliced mRNAs from the nucleus. Proc. Natl. Acad. Sci. U.S.A **98**:1030-1035.

191. **Rosenberg, N. and P. Jolicoeur**. 1997. Retroviral Pathogenesis, p. 475-585. *In* J. M. Coffin, S. H. Hughes, and H. E. Varmus (eds.), Retroviruses. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

192. **Rosqvist, R., A. Forsberg, M. Rimpilainen, T. Bergman, and H. Wolf-Watz**. 1990. The cytotoxic protein YopE of Yersinia obstructs the primary host defence. Mol. Microbiol. **4**:657-667.

193. **Rosqvist, R. and H. Wolf-Watz**. 1986. Virulence plasmid-associated HeLa cell induced cytotoxicity of Yersinia pseudotuberculosis. Microb. Pathog. **1**:229-240.

194. **Ross, S. R., J. J. Schofield, C. J. Farr, and M. Bucan**. 2002. Mouse transferrin receptor 1 is the cell entry receptor for mouse mammary tumor virus. Proc. Natl. Acad. Sci. U. S. A **99**:12386-12390.

195. **Rout, M. P., J. D. Aitchison, A. Suprapto, K. Hjertaas, Y. Zhao, and B. T. Chait**. 2000. The yeast nuclear pore complex: composition, architecture, and transport mechanism. J. Cell Biol. **148**:635-651.

196. **Saavedra, C., B. Felber, and E. Izaurralde**. 1997. The simian retrovirus-1 constitutive transport element, unlike the HIV-1 RRE, uses factors required for cellular mRNA export. Curr. Biol. **7**:619-628.

197. **Sandrin, V., S. J. Russell, and F. L. Cosset**. 2003. Targeting retroviral and lentiviral vectors. Curr. Top. Microbiol. Immunol. **281**:137-178.

198. **Sansonetti, P.** 2001. Phagocytosis of bacterial pathogens: implications in the host response. Semin. Immunol. **13**:381-390.

199. **Sansonetti, P. J.** 2001. Microbes and microbial toxins: paradigms for microbial-mucosal interactions III. Shigellosis: from symptoms to molecular pathogenesis. Am. J. Physiol Gastrointest. Liver Physiol **280**:G319-G323.

200. **Sansonetti, P. J., J. Arondel, J. R. Cantey, M. C. Prevost, and M. Huerre**. 1996. Infection of rabbit Peyer's patches by *Shigella flexneri*: effect of adhesive or invasive bacterial phenotypes on follicle-associated epithelium. Infect. Immun. **64**:2752-2764.

201. **Sansonetti, P. J. and A. Phalipon**. 1999. M cells as ports of entry for enteroinvasive pathogens: mechanisms of interaction, consequences for the disease process. Semin. Immunol. **11**:193-203.

202. **Schultz, A. M. and S. Oroszlan**. 1983. In vivo modification of retroviral gag gene-encoded polyproteins by myristic acid. J. Virol. **46**:355-361.

203. **Schwartz, D. E., R. Tizard, and W. Gilbert**. 1983. Nucleotide sequence of Rous sarcoma virus. Cell **32**:853-869.

204. **Segref, A., K. Sharma, V. Doye, A. Hellwig, J. Huber, R. Luhrmann, and E. Hurt**. 1997. Mex67p, a novel factor for nuclear mRNA export, binds to both poly(A)+ RNA and nuclear pores. EMBO J. **16**:3256-3271.

205. **Shackleford, G. M. and H. E. Varmus**. 1988. Construction of a clonable, infectious, and tumorigenic mouse mammary tumor virus provirus and a derivative genetic vector. Proc. Natl. Acad. Sci. U. S. A **85**:9655-9659.

206. **Siomi, H., H. Shida, S. H. Nam, T. Nosaka, M. Maki, and M. Hatanaka**. 1988. Sequence requirements for nucleolar localization of human T cell leukemia virus type I pX protein, which regulates viral RNA processing. Cell **55**:197-209.

207. **Sizemore, D. R., A. A. Branstrom, and J. C. Sadoff**. 1995. Attenuated *Shigella* as a DNA delivery vehicle for DNA-mediated immunization. Science **270**:299-302.

208. **Skalka, A. M.** 1989. Retroviral proteases: first glimpses at the anatomy of a processing machine. Cell **56**:911-913.

209. **Smith, K. R.** 2003. Gene therapy: theoretical and bioethical concepts. Arch. Med. Res. **34**:247-268.

210. **Sory, M. P., A. Boland, I. Lambermont, and G. R. Cornelis**. 1995. Identification of the YopE and YopH domains required for secretion and internalization into the cytosol of macrophages, using the cyaA gene fusion approach. Proc. Natl. Acad. Sci. U. S. A **92**:11998-12002.

211. **Sory, M. P. and G. R. Cornelis**. 1994. Translocation of a hybrid YopE-adenylate cyclase from *Yersinia enterocolitica* into HeLa cells. Mol. Microbiol. **14**:583-594.

212. **St George, J. A.** 2003. Gene therapy progress and prospects: adenoviral vectors. Gene Ther. **10**:1135-1141.

213. **Staffa, A. and A. Cochrane**. 1994. The tat/rev intron of human immunodeficiency virus type 1 is inefficiently spliced because of suboptimal signals in the 3' splice site. J. Virol. **68**:3071-3079.

214. **Staffa, A. and A. Cochrane**. 1995. Identification of positive and negative splicing regulatory elements within the terminal tat-rev exon of human immunodeficiency virus type 1. Mol. Cell Biol. **15**:4597-4605.

215. **Steinkamp, J. A., B. E. Lehnert, and N. M. Lehnert**. 1999. Discrimination of damaged/dead cells by propidium iodide uptake in immunofluorescently labeled populations analyzed by phase-sensitive flow cytometry. J. Immunol. Methods **226**:59-70.

216. **Stewart, L. and V. M. Vogt**. 1991. trans-acting viral protease is necessary and sufficient for activation of avian leukosis virus reverse transcriptase. J. Virol. **65**:6218-6231.

217. **Stewart, M., R. P. Baker, R. Bayliss, L. Clayton, R. P. Grant, T. Littlewood, and Y. Matsuura**. 2001. Molecular mechanism of translocation through nuclear pore complexes during nuclear protein import. FEBS Lett. **498**:145-149.

218. **Stoltzfus, C. M. and S. J. Fogarty**. 1989. Multiple regions in the Rous sarcoma virus src gene intron act in cis to affect the accumulation of unspliced RNA. J. Virol. **63**:1669-1676.

219. **Strambio-de-Castillia, C. and E. Hunter**. 1992. Mutational analysis of the major homology region of Mason-Pfizer monkey virus by use of saturation mutagenesis. J. Virol. **66**:7021-7032.

220. **Strasser, K. and E. Hurt**. 2000. Yra1p, a conserved nuclear RNA-binding protein, interacts directly with Mex67p and is required for mRNA export. EMBO J. **19**:410-420.

221. **Stutz, F., A. Bachi, T. Doerks, I. C. Braun, B. Seraphin, M. Wilm, P. Bork, and E. Izaurralde**. 2000. REF, an evolutionary conserved family of hnRNP-like proteins, interacts with TAP/Mex67p and participates in mRNA nuclear export. RNA. **6**:638-650.

222. **Stutz, F. and M. Rosbash**. 1998. Nuclear RNA export. Genes Dev. **12**:3303-3319.

223. **Suzuki, T. and C. Sasakawa**. 2001. Molecular basis of the intracellular spreading of Shigella. Infect. Immun. **69**:5959-5966.

224. **Swanson, A. K. and C. M. Stoltzfus**. 1998. Overlapping cis sites used for splicing of HIV-1 env/nef and rev mRNAs. J. Biol. Chem. **273**:34551-34557.

225. **Tabernero, C., A. S. Zolotukhin, J. Bear, R. Schneider, G. Karsenty, and B. K. Felber**. 1997. Identification of an RNA sequence within an intracisternal-A particle element able to replace Rev-mediated posttranscriptional regulation of human immunodeficiency virus type 1. J. Virol. **71**:95-101.

226. **Tabernero, C., A. S. Zolotukhin, A. Valentin, G. N. Pavlakis, and B. K. Felber**. 1996. The posttranscriptional control element of the simian retrovirus type 1 forms an extensive RNA secondary structure necessary for its function. J. Virol. **70**:5998-6011.

227. **Tanaka, H., Y. Dong, Q. Li, S. Okret, and J. A. Gustafsson**. 1991. Identification and characterization of a cis-acting element that interferes with glucocorticoid-inducible activation of the mouse mammary tumor virus promoter. Proc. Natl. Acad. Sci. U. S. A **88**:5393-5397.

228. **Tanese, N. and S. P. Goff**. 1988. Domain structure of the Moloney murine leukemia virus reverse transcriptase: mutational analysis and separate expression of the DNA polymerase and RNase H activities. Proc. Natl. Acad. Sci. U. S. A **85**:1777-1781.

229. **Tang, H., K. L. Kuhen, and F. Wong-Staal**. 1999. Lentivirus replication and regulation. Annu. Rev. Genet. **33**:133-170.

230. **Thiele, H. G.** 1991. Lymphocyte homing: an overview. Immunol. Res. **10**:261-267.

231. **Tiley, L. S., M. H. Malim, H. K. Tewary, P. G. Stockley, and B. R. Cullen**. 1992. Identification of a high-affinity RNA-binding site for the human immunodeficiency virus type 1 Rev protein. Proc. Natl. Acad. Sci. U. S. A **89**:758-762.

232. **Tisne, C., B. P. Roques, and F. Dardel**. 2003. Specific recognition of primer tRNA Lys 3 by HIV-1 nucleocapsid protein: involvement of the zinc fingers and the N-terminal basic extension. Biochimie **85**:557-561.

233. **Traina-Dorge, V. and J. C. Cohen**. 1983. Molecular genetics of mouse mammary tumor virus. Curr. Top. Microbiol. Immunol. **106**:35-56.

234. **Truant, R. and B. R. Cullen**. 1999. The arginine-rich domains present in human immunodeficiency virus type 1 Tat and Rev function as direct importin beta-dependent nuclear localization signals. Mol. Cell Biol. **19**:1210-1217.

235. **Turner, G., M. Barbulescu, M. Su, M. I. Jensen-Seaman, K. K. Kidd, and J. Lenz**. 2001. Insertional polymorphisms of full-length endogenous retroviruses in humans. Curr. Biol. **11**:1531-1535.

236. **van Leeuwen, F. and R. Nusse**. 1995. Oncogene activation and oncogene cooperation in MMTV-induced mouse mammary cancer. Semin. Cancer Biol. **6**:127-133.

237. **van Ooyen, A. J., R. J. Michalides, and R. Nusse**. 1983. Structural analysis of a 1.7-kilobase mouse mammary tumor virus-specific RNA. J. Virol. **46**:362-370.

238. **Vink, C., D. C. van Gent, Y. Elgersma, and R. H. Plasterk**. 1991. Human immunodeficiency virus integrase protein requires a subterminal position of its viral DNA recognition sequence for efficient cleavage. J. Virol. **65**:4636-4644.

239. **Vink, C., E. Yeheskiely, G. A. van der Marel, J. H. van Boom, and R. H. Plasterk**. 1991. Site-specific hydrolysis and alcoholysis of human immunodeficiency virus DNA termini mediated by the viral integrase protein. Nucleic Acids Res. **19**:6691-6698.

240. **Vorburger, S. A. and K. K. Hunt**. 2002. Adenoviral gene therapy. Oncologist. **7**:46-59.

241. **Wang, J. and M. F. Wilkinson**. 2000. Site-directed mutagenesis of large (13-kb) plasmids in a single-PCR procedure. Biotechniques **29**:976-978.

242. **Wassef, J. S., D. F. Keren, and J. L. Mailloux**. 1989. Role of M cells in initial antigen uptake and in ulcer formation in the rabbit intestinal loop model of shigellosis. Infect. Immun. **57**:858-863.

243. **Wei, S. Q., K. Mizuuchi, and R. Craigie**. 1998. Footprints on the viral DNA ends in moloney murine leukemia virus preintegration complexes

reflect a specific association with integrase. Proc. Natl. Acad. Sci. U. S. A **95**:10535-10540.

244. **Weis, K.** 2002. Nucleocytoplasmic transport: cargo trafficking across the border. Curr. Opin. Cell Biol. **14**:328-335.

245. **Weiss, S. and T. Chakraborty**. 2001. Transfer of eukaryotic expression plasmids to mammalian host cells by bacterial carriers. Curr. Opin. Biotechnol. **12**:467-472.

246. **Weiss, S. and S. Krusch**. 2001. Bacteria-mediated transfer of eukaryotic expression plasmids into mammalian host cells. Biol. Chem. **382**:533-541.

247. **Wiedle, G., D. Dunon, and B. A. Imhof**. 2001. Current concepts in lymphocyte homing and recirculation. Crit Rev. Clin. Lab Sci. **38**:1-31.

248. **Wilkins, J. A., D. Stupack, S. Stewart, and S. Caixia**. 1991. Beta 1 integrin-mediated lymphocyte adherence to extracellular matrix is enhanced by phorbol ester treatment. Eur. J. Immunol. **21**:517-522.

249. **Wilson, J. M.** 1997. Vectors--shuttle vehicles for gene therapy. Clin. Exp. Immunol. **107 Suppl 1**:31-32.

250. **Wilusz, C. J., W. Wang, and S. W. Peltz**. 2001. Curbing the nonsense: the activation and regulation of mRNA surveillance. Genes Dev. **15**:2781-2785.

251. **Wodrich, H., J. Bohne, E. Gumz, R. Welker, and H. G. Krausslich**. 2001. A new RNA element located in the coding region of a murine endogenous retrovirus can functionally replace the Rev/Rev-responsive element system in human immunodeficiency virus type 1 Gag expression. J. Virol. **75**:10670-10682.

252. **Wodrich, H., A. Schambach, and H. G. Krausslich**. 2000. Multiple copies of the Mason-Pfizer monkey virus constitutive RNA transport element lead to enhanced HIV-1 Gag expression in a context-dependent manner. Nucleic Acids Res. **28**:901-910.

253. **Wrona, T. J., M. Lozano, A. A. Binhazim, and J. P. Dudley**. 1998. Mutational and functional analysis of the C-terminal region of the C3H mouse mammary tumor virus superantigen. J. Virol. **72**:4746-4755.

254. **Xu, L., T. J. Wrona, and J. P. Dudley**. 1996. Exogenous mouse mammary tumor virus (MMTV) infection induces endogenous MMTV sag expression. Virology **215**:113-123.

255. **Xu, L., T. J. Wrona, and J. P. Dudley**. 1997. Strain-specific expression of spliced MMTV RNAs containing the superantigen gene. Virology **236**:54-65.

256. **Yang, J. and B. R. Cullen**. 1999. Structural and functional analysis of the avian leukemia virus constitutive transport element. RNA. **5**:1645-1655.

257. **Yazdanbakhsh, K., C. G. Park, G. M. Winslow, and Y. Choi**. 1993. Direct evidence for the role of COOH terminus of mouse mammary tumor virus superantigen in determining T cell receptor V beta specificity. J. Exp. Med. **178**:737-741.

258. **Yoshida, H., Y. Tomiyama, J. Ishikawa, K. Oritani, I. Matsumura, M. Shiraga, T. Yokota, Y. Okajima, M. Ogawa, J. Miyagawa, T. Nishiura, and Y. Matsuzawa**. 2000. Integrin-associated protein/CD47 regulates motile activity in human B-cell lines through CDC42. Blood **96**:234-241.

259. **Zhang, W. W.** 1999. Development and application of adenoviral vectors for gene therapy of cancer. Cancer Gene Ther. **6**:113-138.

260. **Zhang, Y., J. C. Rassa, M. E. deObaldia, L. M. Albritton, and S. R. Ross**. 2003. Identification of the receptor binding domain of the mouse mammary tumor virus envelope protein. J. Virol. **77**:10468-10478.

261. **Zhou, Z., M. J. Luo, K. Straesser, J. Katahira, E. Hurt, and R. Reed**. 2000. The protein Aly links pre-messenger-RNA splicing to nuclear export in metazoans. Nature **407**:401-405.

262. **Zhu, Q. and J. P. Dudley**. 2002. CDP binding to multiple sites in the mouse mammary tumor virus long terminal repeat suppresses basal and glucocorticoid-induced transcription. J. Virol. **76**:2168-2179.

263. **Zhu, Q., K. Gregg, M. Lozano, J. Liu, and J. P. Dudley**. 2000. CDP is a repressor of mouse mammary tumor virus expression in the mammary gland. J. Virol. **74**:6348-6357.

264. **Zolotukhin, A. S., A. Valentin, G. N. Pavlakis, and B. K. Felber**. 1994. Continuous propagation of RRE(-) and Rev(-)RRE(-) human immunodeficiency virus type 1 molecular clones containing a cis-acting element of simian retrovirus type 1 in human peripheral blood lymphocytes. J. Virol. **68**:7944-7952.

Vita

Melissa Sue Simper was born in San Antonio, TX on May 14, 1975, the daughter of Herbert R. and Jeannette Mann. After attending Flatonia High School for her high school education, she entered St. Mary's University in the fall of 1993. In May, 1997, she graduated *magna cum laude* with a Bachelor of Science degree in Biology. In the fall of 1997, she entered the Molecular Biology Ph.D. program at The University of Texas at Austin; joining Dr. Jaquelin Dudley's lab in August, 1998.

Permanent address: P.O. Box 342

Moulton, TX 77975

This dissertation was typed by the author.